

MAKING SURE THAT MATHEMATICS EDUCATION
RESEARCH REACHES THE CLASSROOM

PROCEEDINGS

of the 48th Conference of the International Group
for the Psychology of Mathematics Education

JULY 28 TO AUGUST 2, 2025
SANTIAGO, CHILE

EDITORS

Claudia Cornejo, Patricio Felmer, David M. Gómez, Pablo Dartnell,
Paulina Araya, Armando Peri, Valeria Randolph

RESEARCH REPORTS — VOLUME 2
K — Z



Universidad
de Chile

LOH
Universidad
de O'Higgins



48th Conference of the International
Group for the Psychology of
Mathematics Education (PME)

July 28 to August 2, 2025

Santiago, Chile

Website:

<https://eventos.cmm.uchile.cl/pme48>

Cite as:

Cornejo, C., Felmer, P., Gómez, D. M., Dartnell, P., Araya, P., Peri, A., & Randolph, V. (Eds.) (2025). *Proceedings of the 48th Conference of the International Group for the Psychology of Mathematics Education: Research Reports, Vol. 2*. PME.

Proceedings are also available on the IGPME website:

<https://www.igpme.org/>

Copyright © 2025 left to authors

All rights reserved

ISSN 2790-3648

Printed by:

Impresora y Comercial Feyser Ltda., www.feyser.cl

Logo and proceedings cover designed by:

Sebastián Lagunas López, sebalagunasl@gmail.com

TABLE OF CONTENTS
PROCEEDINGS OF PME 48
RESEARCH REPORTS
VOLUME 2 (K – Z)

Table of contents. Proceedings of PME 48: Research Reports, Vol. 2	i
Exploring the relationships between individual cognitive activation, self-efficacy and test-anxiety in upper secondary mathematics Karlsson, Jimmy	3
Exploring collaboration: The effect of gender on mathematics learning preferences Kim, Sang Hyun; Evans, Tanya	11
Proof videos in undergraduate education: viewing strategies and proof comprehension Kirsten, Katharina; Wirth, Laura	19
Theorizing opportunities to learn as a basis for inclusive mathematics education Kress, Nancy Emerson; Duran Oliva, Pablo	27
Adhering to principles for quality teaching: Measuring facilitator assessments of a fictional PD situation Lehmann, Malte; Shure, Victoria; Rothering, Sophia; Laschke, Christin; Roesken-Winter, Bettina	35
‘Mathematical talk videos’: Empirically initiated conceptual considerations Leoci, Lena; Vogel, Rose F.	43
How do parental and teacher support influence high school students' mathematics learning motivation? — a chain mediation model based on growth mindset and self-efficacy Li, Xinxin; Wijaya, Tommy Tanu; Cao, Yiming	51
Secondary students' understanding of the functions of proof for the equivalent meanings of general validity Lundt, Sarah; Damrau, Milena; Ufer, Stefan	59

Embodied argumentation: Convincing yourself, your friend, or the skeptic? Marco, Nadav; Palatnik, Alik	67
I beg to differ: The germ cell of mathematical debate Matteucci, Amedeo; Veronesi, Ilaria; Tortoriello, Francesco Saverio	75
The role of domain-general cognitive skills in early mathematical development: A longitudinal study of 4- and 5-year-old children Mendez Jurado, Maria Agustina; Pearson, Rufina	83
Mathematical self-concept and experience of competence while modelling with and without experiments Menzel, Marielena; Kosiol, Timo; Rach, Stefanie; Geisler, Sebastian	91
Fostering preservice teachers' MWS-interpretative knowledge through epistemological Taxicab perturbations Miranda, Annamaria; Espinoza-Vasquez, Gonzalo; Menares, Romina; Verdugo, Paula	99
Inter-component changes in mathematics teaching activity - focus on potential core change processes Möller, Victoria	107
Didactic and disciplinary knowledge of elementary school teachers in Costa Rica: A study of statistics Morales Reyes, José Luis; Alfaro Arce, Ana Lucía	115
Implications for teacher training of inconsistencies between university and secondary textbooks regarding the multiplication principle Morcos, Jonathan; González-Martín, Alejandro S.	123
Prospective mathematics teachers' perceptions on collaborative lesson study Nagari-Haddif, Galit; Arcavi, Abraham; Karsenty, Ronnie	131
Transition from dynamic to static digital artifacts to promote meta-level learning: A case study with scales and early algebraic discourse Nannini, Bernardo; Antonini, Samuele	139
"Homework should only prepare for the exam!" – students' assessment of (potential) obligatory homework in mathematics education courses at university Neuhaus-Eckhardt, Silke; Siller, Hans-Stefan	147

Exploring the role of internalised standards in comparative judgement in mathematics Palisse, Jennifer; MacLean, Mark; King, Deborah	155
How do cultural norms influence teacher noticing? Paul, Josephine Friederike; Dreher, Anika; Wang, Ting-Ying; Hsieh, Feng-Jui; Lindmeier, Anke	163
Mathematics specialized knowledge of special education teachers to teach numbers and operations to students with visual impairments Piñeiro, Juan Luis; Manolino, Carola; Pinto, Eder; Román, Catalina	171
Analysis of teachers' expertise regarding combinatorics Pöhler-Friedrich, Birte; Thomas, Charlott	179
Mapping routines for analyzing tutor-students interactions Poli, Cristina; Lisarelli, Giulia; Baccaglini-Frank, Anna	187
Analyzing inquiry-based learning with problem posing in mathematics teacher education: Insights from the expansive learning cycle Pozhegu-Ermeni, Gresa; Sikko, Svein Arne; Grimeland, Benedikte	195
Eye movements and spontaneous focusing on numerosity among five-year-old children Rajalakso, Satu Marianne; Södervik, Ilona; Hannula-Sormunen, Minna; Hannula, Markku	203
How students evaluate different task formats? An exploratory study in a large mathematics course for business students Reich, Birte; Liebendörfer, Michael; Eichler, Andreas; Schürmann, Mirko; Stroh-Maraun, Nadja; Müller, Raphael	211
Understanding the solution set of systems of linear equations through geometric visualization Rozas Torres, Elizabeth Noemí; Cárcamo Bahamonde, Andrea; Fortuny Aymemí, Josep Maria	219
Teachers & bold problem-solving: Self-orientation & gendered-orientation Rubel, Laurie H; Shabari, Juhaina A; Saleh, Amal	227
Components of computational thinking demonstrated by preservice mathematics teachers during a math modelling task Santibáñez Galdames, Patricio; Bustos-Osorio, Bárbara; Castillo-Funes, Josefa	235

Prospective secondary mathematics teachers' use of inquiry-based teaching principles to plan a lesson Santonja Serrano, Paloma; Moreno Moreno, Mar; Llinares Ciscar, Salvador	243
It is time: An eye-tracking study on clock-reading Schindler, Maïke; Kölsch, Mia-Milena; Asghari, Parviz; Simon, Anna Lisa	251
Real-world connections in mathematics education for students with intellectual disabilities: A systematic literature review Schmidt, Laura; Janina, Krawitz; Susanne, Schnepel	259
Moving beyond moves: Redefining mathematically responsive teaching as a humanistic stance Schwartz, Gil	267
Pre-service primary teachers' diagnostic judgements: Effect of teacher dispositions and student characteristics Seibold, Tjorven Lea; Meyer-Jenßen, Lars; Laschke, Christin; Roesken-Winter, Bettina	275
Locating numbers on the number line: An eye-tracking study with first graders with and without risk of developing mathematical difficulties Simon, Anna Lisa; Geimer, Laura; Baumanns, Lukas; Pitta-Pantazi, Demetra; Christou, Constantinos; Lilienthal, Achim J.; Schindler, Maïke	283
Teacher leaders' noticing in modelling and argumentation in the mathematics classroom Solar, Horacio; Gómez, Florencia; Martínez, María Victoria; Ortiz, Andrés; Arriagada, Victoria	291
Perceived difficulty in mathematics education Spagnolo, Camilla	299
Using knowledge about proof principles for proof validation at secondary school level Sporn, Femke; Sommerhoff, Daniel; Heinze, Aiso	307
The cultural politics of Korean mathematics curriculum reforms through Asia as method Sum, Emily	315

Examining equitable participation with K-12 teacher leaders Thanheiser, Eva; Roman, Kathryn; Heaton, Ruth	323
Analysis of students' self-written word problems related to arrangement without repetition Thomas, Charlott; Pöhler, Birte	331
The moderating role of self-criticism in the relationship between mathematics confidence and anxiety Ting, Yu-Shan; Seah, Wee Tiong	339
Establishing practices of unpacking meanings of mathematical concepts in multiple languages Uribe, Ángela; Prediger, Susanne	347
Extramathematical connections recognized by prospective elementary school teachers Vargas Herrera, Juan Pablo; Vanegas, Yuly	355
Exploring teachers' attempts aiming struggling and competent students' equitable participation Vasilopoulou, Maria; Triantafillou, Chrissavgi	363
Mathematicians' views on the subject-specific education of prospective teachers Weber, Birke-Johanna; Dreher, Anika; Heinze, Aiso; Sommerhoff, Daniel	371
Ecological conceptualizations in mathematics teacher education and teacher noticing Weston, Tracy L.	379
The influence of mathematical achievement on the development of intuitive probabilistic reasoning in lower secondary school students Wiemann, Leon; Schroeder, Sascha; Halverscheid, Stefan	387
Likes, views, comments: How is viewer engagement related to high- and low-quality explanatory videos? Wölck, Laura; Bednorz, David; Heinze, Aiso	395
Students' self-assessment in the field of fractions: How accurate are they? Wolff, Alexander; Bednorz, David; Schmidt-Thieme, Barbara	403

Noticing of preservice primary school teachers in mathematics while watching own video vignettes Wöller, Susanne	411
Can Germany learn from "Digital Denmark"? Do the digitalization-related attitudes of German and Danish teachers differ? Wulff, Mira Hykkelbjerg; Radkowitsch, Anika; Heinze, Aiso	419
Unexpected richness from a wooden arch puzzle: Embodied ways of learning about parabolas Yu, Paul Woo Dong; Oslund, Joy A.; Alshuli, Tasnim A.	427
Axes without numbers: What function can this be? Zazkis, Rina	435
Urbanization as a catalyst of fraction understanding: A case study Zelbo, Sian E	443



**MAKING SURE THAT MATHEMATICS EDUCATION
RESEARCH REACHES THE CLASSROOM**

**RESEARCH REPORTS
K – Z**

EXPLORING THE RELATIONSHIPS BETWEEN INDIVIDUAL COGNITIVE ACTIVATION, SELF-EFFICACY AND TEST-ANXIETY IN UPPER SECONDARY MATHEMATICS

Jimmy Karlsson

Department of Mathematics, Karlstad University, Sweden

In this exploratory study of 1009 upper secondary students from 49 classrooms, two dimensions of individual cognitive activation were distinguished: task demands and interaction demands. Both dimensions contributed to students' self-efficacy, although task demands ($\beta=0.235$, $p <.001$) to a greater extent than interaction demands ($\beta=0.176$, $p <.001$). Higher task demands predicted increased test-anxiety ($\beta=0.109$, $p <.001$), though the negative indirect effect via self-efficacy ($\beta=-0.087$, $p <.001$) rendered the total effect non-significant. Interaction demands showed an indirect relationship with test-anxiety via self-efficacy ($\beta=-0.065$, $p <.001$). These findings show the importance of students' individual perceived cognitive activation related to self-efficacy and test-anxiety, with potential implications for theory and practice.

INTRODUCTION

Cognitive activation related to mathematics classroom practices, focuses on the depth of students' cognitive engagement by promoting higher order thinking for development of conceptual understanding (Lipowsky et al., 2009; Praetorius et al., 2018). Explored as a classroom construct, cognitive activation has been linked to higher mathematics achievement (Kunter et al., 2013; Praetorius et al., 2018; Spreitzer et al., 2022). Cognitive activation has also more recently been associated with self-related perceptions such as self-concept, self-efficacy, and also mathematics anxiety (Liu et al., 2022; Scherer et al., 2016; Zuo et al., 2024), extending original frameworks that mainly related cognitive activation to achievement and not aspects such as self-efficacy. From an opportunity – use perspective, cognitive activation has mainly been explored as a classroom opportunity (Praetorius et al., 2018) and less is known about individual students' perceived cognitive activation, as in the latter part considering use. Hence, there is a potential to extend the opportunity – use perspective by also exploring the individual student perception of cognitive activation, which complements the perspective by putting emphasis on the latter, students' use, and their individual cognitive engagement. Scherer et al. (2016) suggested that exploring between-student differences could complement research on shared classroom learning environment. The individual student approach could also be extended to provide insights into indirect relationships, connecting individual students perceived cognitive activation and self-related perceptions such as self-efficacy and mathematics test-anxiety.

The aim of this paper is to explore individual perceived cognitive activation and its relationship with self-related perceptions in terms of self-efficacy and mathematics

test-anxiety. To pursue this aim we first examine the construct of individual cognitive activation by a newly developed instrument and then explore this construct together with self-efficacy and mathematics test-anxiety.

Cognitive activation

Potential cognitive activation is a multidimensional construct, including dimensions such as challenging tasks, challenging questions/discussions, activating prior knowledge, and support for metacognition (Praetorius et al., 2018; Spreitzer et al., 2022), with the overall aim of evaluating students cognitive engagement and demands of classroom practices (Lipowsky et al., 2009). Given its multidimensional nature, Lipowsky et al. (2009) raised the question of which dimensions are more central to the construct. Hence, it could be worthwhile to pursue explorations of different dimensions of cognitive activation and how these relate to students' mathematical development.

Cognitive activation has mainly been researched within the framework of Three Basic Dimensions (TBD) of teaching quality. Within the TBD framework, cognitive activation is one of three dimensions, along with classroom management and student support, that has been related to student outcomes in terms of achievement and motivation (Praetorius et al., 2018). It is commonly treated as a shared classroom construct, with instruments focusing on the opportunities provided to students (Praetorius et al., 2018). However, Scherer et al. (2016) studied individual perceptions of instructional quality and identified significant relationships between cognitive activation and achievement and self-related perceptions in terms of self-concept and motivation. Further, Scherer et al. (2016) argue that researching individual perceptions is worthwhile as it could contribute to an understanding of between-students differences, as an alternative approach of studying characteristics of classrooms. However, this requires instruments that instead of having the referent as the mathematics classroom, focus on individual students' perception of cognitive activation.

Self-efficacy, mathematics anxiety and cognitive activation

Self-efficacy refers to a person's self-related perception about their capability to accomplish a future task and is central in social cognitive theory developed by Bandura (1997). It is considered domain-specific and its main sources of development include mastery experience, vicarious experience, verbal persuasion and physiological reactions (Bandura, 1997). Cognitive activation has been associated with self-efficacy (Liu et al., 2022; Zuo et al., 2024), mainly explained by connecting students work with challenging tasks and mastery experiences. Higher self-efficacy has also been linked to higher achievement, effort, and decreased anxiety (Street et al., 2024). However, the relationship between cognitive activation and anxiety has received mixed results where, for example, Kunter et al. (2013) found that classroom-level cognitive activation was not related to achievement anxiety for secondary students in Germany, whereas Liu et al. (2022), and Zuo et al. (2024) identified a significant relationship

where cognitive activation was associated with decreased anxiety in a Chinese middle school students and 14-15 year old students.

Conceptual framework

The constructs and relationships in our conceptual framework include individual student cognitive activation, which stems from the TBD-framework, elaborated to include self-efficacy from social cognitive theory and also extended to mathematics test-anxiety. This connects cognitive engagement from cognitive activation with self-related perceptions, in this study, self-efficacy and also mathematics test-anxiety. Further, the structural relations include both direct and indirect relationships from cognitive activation to self-efficacy and mathematics test-anxiety.

METHOD

Sample and data generation

The sample consists of upper secondary students ($n = 1009$) from 49 different classes in Sweden studying their second mathematics course. Each course commonly involves 70-90 hours of teaching. The sample includes students from both social and science tracks, resulting in a diverse student sample. Students at social tracks most commonly study the second course during their second year, whereas at science-oriented tracks they study more mathematics and take their second course during the latter part of their first year at upper secondary school. Data were generated using a questionnaire which was administered during the last month before the students completed the second course.

Measures

The *cognitive activation* items were developed to assess central elements previously identified. Main difference compared to previously used instruments aiming to measure cognitive activation is the change of referent. Previous instruments commonly framed questions with the teacher or classroom as the referent, such as “Our mathematics teacher expects us too...” or “in our mathematics classroom we are encouraged too...”. In this study, the referent is the individual student’s perception, see items in table 1. All questions were framed as the individual students work during mathematics lessons with the answer options ranging from; (1) rarely or never, (2) sometimes, (3) often, and (4) always or almost always.

Item	Proportion			
	(1)	(2)	(3)	(4)
1 I work with tasks where I have to think thoroughly.	.01	.19	.52	.26
2 I need to think about how different concepts are related.	.05	.36	.43	.16
3 I work with tasks that challenge me.	.03	.19	.46	.31
4 I justify my solutions to tasks.	.07	.34	.39	.20

5	I can keep up in discussions.	.06	.28	.43	.22
6	I need to understand how others think.	.10	.52	.31	.07
7	I need to explain how I think.	.07	.37	.41	.14
8	I need to give reasons for a solution strategy.	.22	.42	.28	.08

Table 1: Cognitive activation items with student answer proportions.

The *mathematics self-efficacy* items were developed to measure students perceived competence to solve tasks that was central to their second mathematics course content. The question stem was: How confident are you regarding the following tasks? Answer options was; (1) Not at all confident, (2) not confident, (3) confident (4) very confident. All items are given in table 2.

Item		Proportion			
		(1)	(2)	(3)	(4)
1	Draw a graph, by hand, to the function described by: $y = x^2 + 4x$.	.15	.36	.31	.17
2	Solve an equation of the type: $x^2 - 8x + 15 = 0$.	.03	.06	.26	.64
3	Explain when a quadratic equation has one, two or no real solutions.	.10	.27	.36	.25
4	Determine the axis of symmetry for a quadratic function.	.06	.15	.35	.42
5	Solve a system of linear equations with an algebraic method.	.05	.17	.38	.39
6	Determine the highest point for a thrown ball which curve is described by a quadratic function.	.10	.25	.33	.31
7	Determine the values for x which satisfies $(x - 4)(x - 5) = 0$.	.03	.10	.28	.58
8	Explain how zeros to a quadratic function can be determined.	.05	.19	.36	.38

Table 2: Self-efficacy items with student answer proportions.

Mathematics test-anxiety was approached by an instrument focusing on preparing for and taking tests in mathematics adapted from the Global Teaching InSight project (OECD, 2020). It contained three items where students answered to what extent the statements would apply to them. Example item: *I get very tense when I study for a mathematics test*. The answer options were ranging from; (1) matches poorly, (2) matches quite poorly, (3) matches quite well, (4) matches very well. Used as a mean score for each student (mean = 2.64, sd = 0.84) with a Cronbach α of 0.81.

Analytic approach

The first step in the analysis includes confirmatory factor analysis (CFA) of the constructs of cognitive activation and mathematics self-efficacy. Then the structural relations were added by exploring how self-efficacy and test-anxiety can be predicted from cognitive activation by means of structural equation modelling (SEM). As a final step in the analysis, the indirect relationships were explored by a 10 000-draw bootstrap

approach. Data analysis was conducted in Mplus ver 8.5 (Muthén & Muthén, 1998-2017). Global fit was evaluated by χ^2 goodness-of-fit test, RMSEA, CFI and SRMR. Due to the ordered categorical nature of cognitive activation and self-efficacy constructs items, the WLSMV estimator was used, which is a robust weighted least squares estimator (Muthén & Muthén, 1998-2017). Of students completing the questionnaire, at most 1.2 % of data was missing for a single item. Therefore, the amount of missing data was not deemed substantial and further, missing at random was assumed, for which the chosen estimator has been shown to provide consistent estimates. During data screening, it was found that some items (e.g., *Determine the axis of symmetry for a quadratic function*) had an intra-class correlation (ICC) of above 0.05, giving support of a clustering effect due to students being nested in groups. To take clustering into account, the complex option in Mplus was used, which compensates for the clustering effect, giving an increase in standard errors but not affecting the estimates.

Preliminary analysis

Using all cognitive activation items in a CFA gave a model which did not fit the data well, $\chi^2(20) = 631.297$, $p < .001$ with RMSEA = .175, CFI = .818 and SRMR = .094. Lowest factor loading was 0.526 with all items significant.

However, separating cognitive activation into two different factors, with item 1-4 constituting CogAct₁ and items 5-8 to CogAct₂ a different fit was obtained. This separation was done based on prior identified dimensions (e.g., challenging tasks and classroom discussions) in the literature of cognitive activation where the CogAct₁ construct represents items related task demands and CogAct₂ represents items related to interaction demands in mathematics. This new model had a fit to the data of $\chi^2(19) = 166.541$, $p < .001$ with RMSEA = .088, CFI = .956 and SRMR = .046 which showed an improved global fit to the data. All loading were above 0.523 and significant. Based on this CFA, a two-factor model was retained. Hence, we continue by using cognitive activation as two factors in our models.

For the Self-Efficacy construct, the CFA gave a model with acceptable fit to the data for a unidimensional construct. $\chi^2(20) = 145.050$, $p < .001$ with RMSEA = .079, CFI = .956 and SRMR .036. All items loaded significantly on the Self-Efficacy factor with loading greater than 0.670.

The analysis then estimated the measurement model including CogAct₁, CogAct₂ and Self-Efficacy. This model did not converge. As the complex option was used and including 49 groups of students, the model was overparameterised with too many free parameters (Kline, 2016). Instead of excluding the complex option from Mplus or using parceling, we reduced the number of items for each factor. Based on loadings and item construction, item 1,2 and 4 was retained for CogAct₁, and for CogAct₂ item 6, 7 and 8. For Self-Efficacy, four items were chosen: 2, 4, 6 and 8. This adjust the

model to include 43 free parameters. This new model showed acceptable fit with $\chi^2(32) = 147.138$, $p < .001$ together with RMSEA = .060, CFI = .973 and SRMR = .039.

RESULTS

The results for the structural model, where we include CogAct₁, CogAct₂, Self-Efficacy and Test-Anxiety are presented in figure 1. This model showed acceptable fit with $\chi^2(39) = 165.166$, $p < .001$ together with RMSEA = .057, CFI = .970 and SRMR = .039. All relationships in the model were found significant, except for the direct relationship between CogAct₂ and Test-Anxiety. For CogAct₁, the model estimated a positive relationship with both Self-Efficacy and Test-Anxiety, meaning that students experiencing higher task demands are also expected to show higher self-efficacy and test-anxiety. Both dimensions of individual cognitive activation in CogAct₁ and CogAct₂ predicts higher Self-Efficacy, although CogAct₁ to a larger extent, based on the standardised estimates.

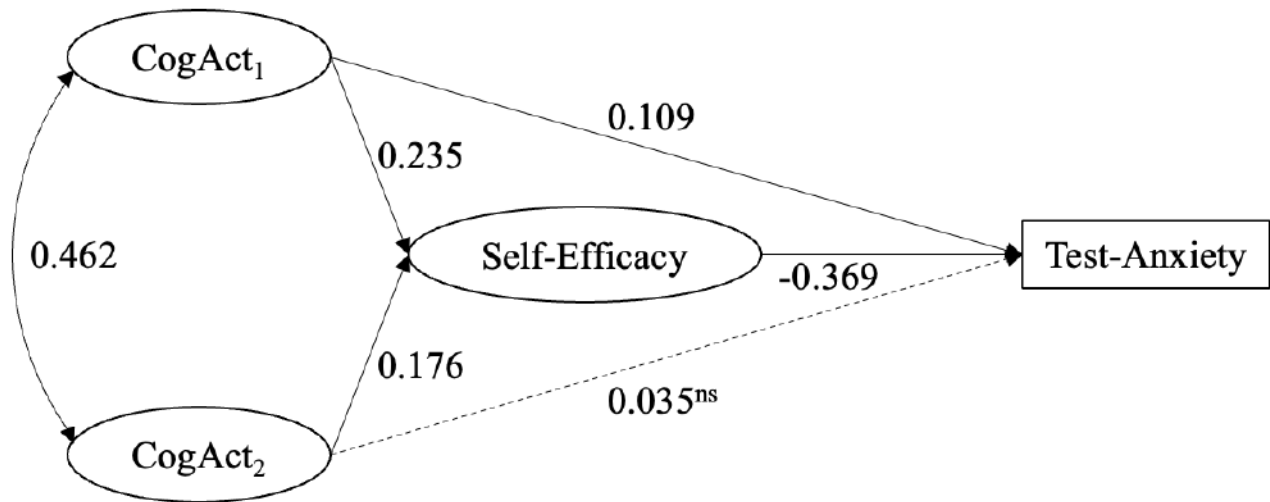


Figure 1: SEM model of cognitive activation, self-efficacy and test-anxiety. Non-significant paths are shown with a dashed line. Standardised estimates.

Finally, both the direct and indirect relationships between CogAct₁, CogAct₂ together with Self-efficacy and Test-Anxiety was estimated. From table 3, we note that although CogAct₁ had a significant direct relationship with Test-Anxiety, the indirect relationship through Self-Efficacy was also significant, rendering the total effect not significantly different from zero (95% CI: -0.083, 0.125). In contrast, CogAct₂ showed a significant indirect (95% CI: -0.107, -0.026), but not direct, relationship with Test-Anxiety.

	Path	Estimate	95% CI
Direct	CogAct ₁ → Anx	0.109	0.016, 0.213
Direct	CogAct ₁ → Self-Efficacy	0.235	0.136, 0.348
Indirect	CogAct ₁ → Self-Efficacy → Test-Anx	-0.087	-0.139, -0.052

Direct	CogAct ₂ → Anx	0.035	-0.055, 0.133
Direct	CogAct ₂ → Self-Efficacy	0.176	0.064, 0.257
Indirect	CogAct ₂ → Self-Efficacy → Test-Anx	-0.065	-0.107, -0.026

Table 3: Direct and indirect relationships between cognitive activation, self-efficacy and mathematics test-anxiety. Standardised estimates with 95% confidence interval.

DISCUSSION AND CONCLUSION

This study extends previous research regarding cognitive activation by focusing on individual perceptions of cognitive engagement and relating this to perceived self-efficacy and mathematical test-anxiety. By means of a new instrument, changing referent from the teacher or classroom to the individual student, the construct of individual perceived cognitive activation was identified through two factors: one relating to the task demands and the other to interaction demands, with different relations to self-efficacy and test-anxiety. Similar subdimensions have also been identified in previous research (Kunter et al., 2013; Lipowsky et al., 2009; Praetorius et al., 2018; Spreitzer et al., 2022) focusing on cognitive activation as a classroom construct. Both factors were significantly related to self-efficacy, which was in line with previous findings by Liu et al. (2022), and Zuo et al. (2024).

Task demands showed a higher estimate ($\beta = 0.235$) compared to interaction demands ($\beta = 0.176$) in the relationships with self-efficacy. This finding underlines the importance of cognitively demanding tasks in mathematics classrooms for student's self-efficacy. This echoes the call by Lipowsky (2009) requesting research exploring the centrality of different dimensions of cognitive activation.

For mathematics test-anxiety, the results showed that both dimensions of individual cognitive activation were indirectly negatively related to mathematics test-anxiety through self-efficacy. As the dimension of mathematical task demands direct relationship with test-anxiety was found to be positively associated with test-anxiety, these results could provide insights related to previous results comparing Kunter et al. (2013) and Zuo et al. (2024). Though it should be noted that previous studies have used different instruments for mathematics related anxiety, and this study focused on mathematics test-anxiety, which could be different from general mathematics anxiety.

This study was based on a cross-sectional design, which limits causal interpretations. Although structural equation modelling allows for testing theoretical relationships (Kline, 2016), further research is needed to replicate the findings, for example by longitudinal or experimental designs, also including other student populations.

In sum, this study's results provide further support for the theoretical extension of frameworks including cognitive activation by also considering relationships with self-related perceptions. For mathematics teachers, the study highlights the importance of engaging students in classroom practices that offer rich experiences with challenging tasks that also encourage explanations and justifications.

REFERENCES

- Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman.
- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (Fourth edition). The Guilford Press.
- Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S., & Neubrand, M. (Eds.). (2013). *Cognitive Activation in the Mathematics Classroom and Professional Competence of Teachers: Results from the COACTIV Project*. Springer US. <https://doi.org/10.1007/978-1-4614-5149-5>
- Lipowsky, F., Rakoczy, K., Pauli, C., Drollinger-Vetter, B., Klieme, E., & Reusser, K. (2009). Quality of geometry instruction and its short-term impact on students' understanding of the Pythagorean Theorem. *Learning and Instruction*, 19(6), 527–537. <https://doi.org/10.1016/j.learninstruc.2008.11.001>
- Liu, Y., Wang, C., Liu, J., & Liu, H. (2022). The role of cognitive activation in predicting mathematics self-efficacy and anxiety among internal migrant and local children. *Educational Psychology*, 42(1), 83–107. <https://doi.org/10.1080/01443410.2021.1987388>
- Muthén, L. K., & Muthén, B. O. (1998). *Mplus User's Guide* (Eighth edition). Muthén & Muthén.
- OECD. (2020). *Global Teaching InSights: A Video Study of Teaching*. OECD. <https://doi.org/10.1787/20d6f36b-en>
- Praetorius, A.-K., Klieme, E., Herbert, B., & Pinger, P. (2018). Generic dimensions of teaching quality: The German framework of Three Basic Dimensions. *ZDM*, 50(3), 407–426. <https://doi.org/10.1007/s11858-018-0918-4>
- Scherer, R., Nilsen, T., & Jansen, M. (2016). Evaluating Individual Students' Perceptions of Instructional Quality: An Investigation of their Factor Structure, Measurement Invariance, and Relations to Educational Outcomes. *Frontiers in Psychology*, 7. <https://doi.org/10.3389/fpsyg.2016.00110>
- Spreitzer, C., Hafner, S., Krainer, K., & Vohns, A. (2022). Effects of Generic and Subject-Didactic Teaching Characteristics on Student Performance in Mathematics in Secondary School: A Scoping Review. *European Journal of Educational Research*, 11(2), 711–737. <https://doi.org/10.12973/eu-jer.11.2.711>
- Street, K. E. S., Malmberg, L.-E., & Schukajlow, S. (2024). Students' mathematics self-efficacy: A scoping review. *ZDM – Mathematics Education*, 56(2), 265–280. <https://doi.org/10.1007/s11858-024-01548-0>
- Zuo, S., Huang, Q., & Qi, C. (2024). The relationship between cognitive activation and mathematics achievement: Mediating roles of self-efficacy and mathematics anxiety. *Current Psychology*. <https://doi.org/10.1007/s12144-024-06700-3>

EXPLORING COLLABORATION: THE EFFECT OF GENDER ON MATHEMATICS LEARNING PREFERENCES

Sang Hyun Kim and Tanya Evans

University of Auckland

This study examines the influence of gender on students' collaborative preferences for learning mathematics (CPLM) over time in an undergraduate mathematics context. Data collected at three points during the semester were analyzed using a two-way mixed ANOVA. Results showed no significant interaction between gender and time, nor a main effect of gender, indicating stable CPLM scores and comparable preferences between male and female students. These findings challenge prior research on gendered differences in collaboration and suggest that further exploration of contextual factors is needed to deepen understanding of CPLM in undergraduate mathematics education.

INTRODUCTION

Amid the growing interest in improving student engagement within undergraduate mathematics education, understanding the diverse preferences that students bring to their learning has become increasingly important. These preferences, particularly in collaborative settings, are shaped by a complex interplay of cognitive, affective, and social factors, and it seems likely that gender may emerge as a salient factor that affects them. The persistent gender differences reported in the literature may suggest that gender can shape students' preferences to collaborate with their peers in mathematics learning contexts (e.g., Else-Quest et al., 2010; Seegers & Boekaerts, 1996). Exploring these dynamics can provide valuable insights into whether gender differences are present and how they may influence student preferences and behaviors within an undergraduate mathematics context. This study seeks to contribute to this area of research by examining the role of gender in shaping collaborative preferences for learning mathematics.

The importance of understanding such collaborative preferences of students is that it may provide us with a greater understanding of how the dispositions held by an individual may be linked to other related social-cognitive and affective factors. In certain contexts, stronger preferences to engage with mathematics in a collaborative manner are likely to drive behaviors within and beyond the classroom that align with them, which could impact learning outcomes, including achievement. A preference for collaboration may foster active participation and idea sharing, promoting social learning processes that enhance conceptual understanding. Alternatively, preferences for independent and individualized work might encourage one to work at a more appropriate pace and develop greater autonomy while problem-solving. Furthermore, these preferences may manifest in contexts beyond academic settings, influencing how

students engage with others and apply mathematical reasoning to real-world problems. In this way, preferences act as a lens through which we can better understand the interplay between students' social tendencies, affective experiences, and their practical use of mathematical knowledge.

Preferences in mathematics

Student preferences in mathematics education reflect individual inclinations toward specific modes of engagement, such as collaborative or independent work. Preferences are an example of a motivational structure that serves as "templates for whether to put forth effort towards mathematical activity and the extent to which efforts are seen as efficacious" (Goldin et al., 2016, p. 18). These preferences are closely tied to students' values and can help shape the tasks they engage with and the pedagogical approach in which they do so. Beliefs are typically grounded in students' perceptions of truth or correctness (Beswick, 2007). On the other hand, preferences are shaped by personal experiences and interests, influencing how students engage with mathematical tasks in various contexts. While preferences can guide short-term behaviors, such as the choice to work with others or tackle a problem independently, they can also evolve into more stable patterns or habits over time, influencing long-term academic engagement (Goldin et al., 2016). These stable preferences can impact not only students' mathematical achievement but also their persistence and development of mathematical identities. Hence, understanding the role of preferences may be important for informing teaching practices that align with students' inclinations and fostering more effective learning experiences.

Collaborative preferences

A desire to work more collaboratively or individually is a natural point of variation worth exploring, with limited studies exploring this at the undergraduate level or within the domain of mathematics. The literature on student experiences during collaboration underscores a strong need to understand ways in which we can better capture the preferences and attitudes of students towards collaboration within mathematics. Existing studies in the broader educational literature have directed attention to some factors that may be correlated with certain preferences. Gajderowicz et al. (2023) found that students preferred working in groups over working individually, although these preferences for group work depended on factors such as subject domain and ability. Furthermore, the authors concluded that understanding student preferences can improve student satisfaction in learning mathematics, noting that "students can enjoy learning math if comfortable study modes are used" (Gajderowicz et al., 2023, pp. 8-9), a conclusion supported by earlier STEM education research (Okebukola, 1986). In fact, Okebukola (1986) found that alignment between students' preferred modes of learning and the mode in which they were taught resulted in the best achievement outcomes. Similarly, aligning teaching practices with student preferences for engagement can result in improvements in conceptual learning and perceptions of autonomy, as demonstrated by Jang et al. (2016). However, some scholars warn against

simply following the desires of the students, as while it could create an illusion of learning, this may not always be reflective of an improvement in student achievement (Deslauriers et al., 2019).

Gender differences

Research on gender differences in mathematics achievement has shown mixed findings, with boys historically outperforming girls in certain contexts, particularly at higher levels of mathematics and in specific regions (Else-Quest et al., 2010; Guiso et al., 2008; Seegers & Boekaerts, 1996); however, this gap has diminished or disappeared in many countries over recent decades (Hyde & Mertz, 2009). Some differences between males and females in affective and motivational factors and their effects on mathematics achievement, including self-concept, attribution of success, and perceptions of difficulty, have also been reported (Ethington, 1992; Mejía-Rodríguez et al., 2021; Wolleat et al., 1980). Research suggests that girls often prefer collaborative environments emphasizing collective reasoning and valuing relationships, while boys may be more inclined to work individually (Gajderowicz et al., 2023; Kanevsky et al., 2022; Owens, 1985). Borgonovi et al. (2023) further contribute to this understanding by emphasizing that these gendered preferences are not solely reflective of individual inclinations but rather are influenced by a range of cognitive, emotional, and social factors. Using 2015 PISA data, the authors reported that girls scored higher than boys in collaborative problem-solving skills across all countries, with this difference being more pronounced in favour of girls in countries with greater gender equality (Borgonovi et al., 2023). These findings highlight how societal contexts and cultural norms shape not only achievement gaps but also students' preferences and approaches to learning. This insight underscores the need to consider gender as a dynamic and influential factor in shaping students' collaborative preferences for learning mathematics within the context of undergraduate education while recognizing that collaboration and student engagement may evolve over time.

RESEARCH AIM

The aim of this study is to explore the influence of gender on students' collaborative preferences for learning mathematics (CPLM) over time. This research addresses two primary questions: (1) *what are the collaborative preferences for learning mathematics among undergraduate students across different genders*, and (2) *how does gender (male vs. female) influence students' collaborative preferences for learning mathematics (CPLM) across different time points?*

METHODOLOGY

Context

This study draws on data gathered during the second semester of 2024 from a second-year undergraduate service mathematics course at a large university in New Zealand. This course introduces students to key mathematical concepts required for various academic programs across the university, covering three main topics: calculus II, linear

algebra II, and differential equations. Delivered over a 12-week semester, the course included weekly one-hour problem-solving sessions (tutorials) in which students were given a set of mathematics problems and were encouraged to work with peers, although this was not mandated. A tutor (either the lecturer or an experienced graduate student) would be available to assist students with problem-solving by providing guidance, clarifying concepts, and answering questions as needed.

Data

The self-report data used in this study was part of a larger project, with data collected at three points in the semester (the first week, after the midsemester break, and the final week). Here, we focus on students' collaborative preferences for learning mathematics (CPLM) and gender. Of the 294 students enrolled in the course, 201 completed the surveys at all three time points. Students who selected 'declined to answer' about their gender ($n = 3$) were not included in the analysis, resulting in a sample size of 198 students for the analysis.

Students' collaborative preferences for learning mathematics were assessed using a 5-item scale developed and validated by Kim and Evans (2025). Students responded to five slider-based, close-ended items, rating their preferences on a scale from 0 (indicating a preference for individual learning) to 100 (indicating a preference for collaborative learning). These ratings captured students' collaborative preferences across various scenarios. Two example items from the scale are *"What is the most effective way for you to learn mathematics?"* and *"In what social setting do you prefer to be exposed to novel concepts?"* To assess students' collaborative preferences for learning mathematics (CPLM), we calculated a composite score for each student at each time point by averaging their responses across the five items.

Analysis

Within this study, we used descriptive statistics to examine students' CPLM based on gender. We also conducted a two-way mixed ANOVA to investigate the effect of gender as a between-subject factor on CPLM scores over time measured at three time points. Gender was the between-subjects factor (male and female) and time was the within-subjects factor (3 levels), with an interaction term to assess whether the change in CPLM scores over time differed by gender.

RESULTS

Table 1 shows the descriptive statistics for the CPLM scores for males and females at the three time points. As shown in Figure 1, the CPLM scores for males were slightly higher than those for females at each time point, suggesting a higher preference for collaboration. The differences between the means of males and females also did not differ significantly across the semester, with the differences between male and female students' CPLM scores at the three time points being 4.44, 3.16, and 5.23, respectively.

CPLM	Gender	Mean	Std. Deviation	N
Time 1	Male	50.40	22.34	115
	Female	45.96	22.28	83
	Total	48.54	22.37	198
Time 2	Male	49.19	23.13	115
	Female	46.03	21.73	83
	Total	47.86	22.55	198
Time 3	Male	52.19	23.73	115
	Female	46.96	21.94	83
	Total	50.00	23.08	198

Table 1: Descriptive statistics for CPLM data by gender across three time points during the semester.

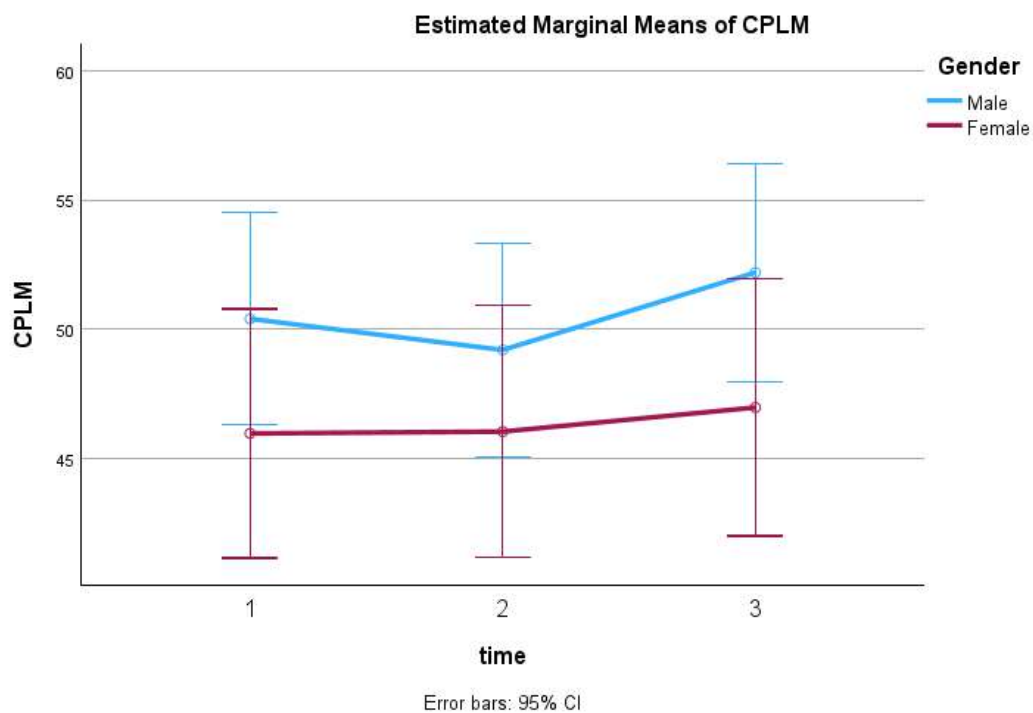


Figure 1: Change in CPLM scores over time for males and females.

Mauchly's test of sphericity revealed a violation of the sphericity assumption for the two-way interaction, $\chi^2(2) = 12.89$, $p = .002$. Greenhouse-Geisser corrections were applied to adjust the degrees of freedom for the F -tests going forward with the analysis of within-subject effects in order to account for this violation ($\epsilon = .904$). No statistically significant interaction between gender and time on CPLM was found, $F(1.880, 368.429) = .260$, $p = .758$, partial $\eta^2 = .001$. The main effect of time did not show a

statistically significant difference in mean CPLM at the different time points, $F(1.880, 368.429) = .976, p = .373$, partial $\eta^2 = .005$.

Similarly, the main effect of gender did not show a statistically significant difference in mean CPLM between males and females, $F(1, 196) = 2.340, p = .128$, partial $\eta^2 = .012$. While these differences were not statistically significant ($p = .128$), the partial eta-squared value ($\eta^2 = .012$) suggests a small effect size, indicating the potential for a gender-related difference. This may warrant further investigation, ideally with a larger sample or over an extended time frame.

DISCUSSION

This study investigated collaborative preferences for learning mathematics (CPLM) within a sample of undergraduate students and whether gender influenced these preferences over time. The mean scores for males were higher than for females at all time points, with the difference between them largest by the end of the semester (5.23). While this suggests males exhibited a slightly higher preference for collaborative learning than females, contrary to prior research findings, these differences were not statistically significant. Notably, the set of 95% confidence intervals of mean values across genders and time points, ranging from 41.1 to 56.4, suggest that most students maintained a balanced preference for both collaborative and individual learning.

Interestingly, the existing literature typically associates males with a preference for individual work (e.g., Gajderowicz et al., 2023; Owens, 1985). Such findings tend to emphasize the social and emotional aspects of collaborative learning, which are traditionally seen as aligning more closely with gender norms favoring girls. However, our results challenge the existing literature, as gender differences in CPLM were not only statistically non-significant but also reversed, with boys preferring more collaboration than girls. This trend was consistent across all three time points during the semester. This could reflect contextual factors specific to this study, such as the structure of the undergraduate tutorials, where collaboration was encouraged but not essential, and where marks were awarded for participation regardless of the work done individually or collaboratively, potentially mitigating any effect on student preferences. Another possible explanation, from a psychological perspective, is the potential benefits that male students may derive from this type of tutorial environment. Raabe and Block (2024) suggest that males may derive social validation and confidence from their performance in mathematics, which is reinforced through peer interactions and friendship networks, whereas females do not feel the same social pressure.

Lastly, no significant interaction effect between gender and time was observed in this study, nor was the main effect of gender found on collaborative preferences for learning mathematics (CPLM). These findings suggest that, within this sample, male and female students exhibited more comparable preferences for collaboration than expected. Moreover, the absence of the main effect of time supports the possibility that CPLM preferences are relatively stable, indicating it may possess a trait-like quality. This

stability could stem from broader and more diverse learning experiences undergraduate students have encountered, both in mathematics education and education more generally, compared to younger students. Regardless, the underlying mechanisms are yet to be understood, warranting further research to examine how transient these preferences are and the influence of social factors around them.

This study has several limitations that future research could address. The small, context-specific sample may limit the generalizability of these findings, emphasizing the need for larger, more diverse samples to better assess CPLM stability and detect subtle gender differences with greater statistical power. In particular, the absence of significant gender differences in this study could be due to the sample size, suggesting that a larger sample may be needed in future studies to reduce the likelihood of a Type II error. Additionally, the study did not consider factors such as cultural background, prior collaborative experiences, or academic achievement, all of which could influence CPLM. Future research could address these gaps by incorporating diverse samples and examining additional factors. Furthermore, integrating observational or interview data could provide deeper insights into how students work and how their preferences develop or change in response to specific events, tasks, or interventions.

In conclusion, this research sought to contribute to our understanding of how collaborative preferences and gender intersect in undergraduate mathematics. Future research can build upon this work by continuing to examine the underlying mechanisms behind these preferences, further uncovering the factors that shape collaborative preferences and their relationship with gender in educational contexts.

References

- Beswick, K. (2007). Teachers' Beliefs That Matter in Secondary Mathematics Classrooms. *Educational Studies in Mathematics*, 65(1), 95-120. <https://doi.org/10.1007/s10649-006-9035-3>
- Borgonovi, F., Han, S. W., & Greiff, S. (2023). Gender differences in collaborative problem-solving skills in a cross-country perspective. *Journal of educational psychology*, 115(5), 747. <https://doi.org/10.1037/edu0000788>
- Deslauriers, L., McCarty, L. S., Miller, K., Callaghan, K., & Kestin, G. (2019). Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. *Proceedings of the National Academy of Sciences*, 116(39), 19251-19257. <https://doi.org/10.1073/pnas.1821936116>
- Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: a meta-analysis. *Psychological Bulletin*, 136(1), 103.
- Ethington, C. A. (1992). Gender Differences in a Psychological Model of Mathematics Achievement. *Journal for research in mathematics education*, 23(2), 166-181. <https://doi.org/10.2307/749499>
- Gajderowicz, T., Jakubowski, M., Wrona, S., & Alkhadim, G. (2023). Is students' teamwork a dreamwork? A new DCE-based multidimensional approach to preferences towards group

- work. *Humanities and Social Sciences Communications*, 10(1), 154. <https://doi.org/10.1057/s41599-023-01641-x>
- Goldin, G. A., Hannula, M. S., Heyd-Metzuyanim, E., Jansen, A., Kaasila, R., Lutovac, S., Di Martino, P., Morselli, F., Middleton, J. A., & Pantziara, M. (2016). Attitudes, beliefs, motivation and identity in mathematics education: An overview of the field and future directions. Springer.
- Guiso, L., Monte, F., Sapienza, P., & Zingales, L. (2008). Culture, Gender, and Math. *Science*, 320(5880), 1164-1165. <https://doi.org/10.1126/science.1154094>
- Hyde, J. S., & Mertz, J. E. (2009). Gender, culture, and mathematics performance. *Proceedings of the National Academy of Sciences*, 106(22), 8801-8807. <https://doi.org/doi:10.1073/pnas.0901265106>
- Jang, H., Reeve, J., & Halusic, M. (2016). A New Autonomy-Supportive Way of Teaching That Increases Conceptual Learning: Teaching in Students' Preferred Ways. *The Journal of Experimental Education*, 84(4), 686-701. <https://doi.org/10.1080/00220973.2015.1083522>
- Kanevsky, L., Lo, C. O., & Marghelis, V. (2022). Individual or collaborative projects? Considerations influencing the preferences of students with high reasoning ability and others their age. *High Ability Studies*, 33(1), 87-119. <https://doi.org/10.1080/13598139.2021.1903842>
- Kim, S. H., & Evans, T. (2025). Collaborative Preferences for Learning Mathematics: A Scale Validation Study. *Proceedings of the 26th Annual Conference on Research in Undergraduate Mathematics Education* [in press],
- Mejía-Rodríguez, A. M., Luyten, H., & Meelissen, M. R. M. (2021). Gender Differences in Mathematics Self-concept Across the World: an Exploration of Student and Parent Data of TIMSS 2015. *International Journal of Science and Mathematics Education*, 19(6), 1229-1250. <https://doi.org/10.1007/s10763-020-10100-x>
- Okebukola, P. A. (1986). The influence of preferred learning styles on cooperative learning in science. *Science education* (Salem, Mass.), 70(5), 509-517. <https://doi.org/10.1002/sce.3730700505>
- Owens, L. (1985). The learning preferences of students and teachers: An Australian-American comparison. *Teaching and Teacher Education*, 1(3), 229-241. [https://doi.org/10.1016/0742-051X\(85\)90006-X](https://doi.org/10.1016/0742-051X(85)90006-X)
- Raabe, I. J., & Block, P. (2024). The gendered maths confidence gap, social influence and social integration. *European Societies*, 26(5), 1596-1631. <https://doi.org/10.1080/14616696.2024.2349217>
- Seegers, G., & Boekaerts, M. (1996). Gender-related differences in self-referenced cognitions in relation to mathematics. *Journal for research in mathematics education*, 27(2), 215-240. <https://doi.org/10.2307/749601>
- Wolleat, P. L., Pedro, J. D., Becker, A. D., & Fennema, E. (1980). Sex Differences in High School Students' Causal Attributions of Performance in Mathematics. *Journal for research in mathematics education*, 11(5), 356-366. <https://doi.org/10.2307/748626>

PROOF VIDEOS IN UNDERGRADUATE EDUCATION: VIEWING STRATEGIES AND PROOF COMPREHENSION

Katharina Kirsten and Laura Wirth

University of Münster, Germany

Despite the increasing use of instructional videos in higher education, our understanding of how to use them beneficially is still limited. This is particularly true for videos that present proofs. In this study, we use the ICAP framework to investigate students' use of different viewing strategies, namely pausing, rewinding, and taking notes, and its relationship to proof comprehension. Correlation analyses based on data from 167 first-year students showed no significant relationship between the use of viewing strategies and students' scores on a proof comprehension test. However, additional qualitative analyses from interviews with seven students suggest that even students with low proof comprehension are actively engaged. Possible explanations and implications are discussed in the light of research on proof comprehension.

INTRODUCTION

In recent years, there has been an overall trend to integrate instructional videos into higher education. In the context of flipped classrooms, instructional videos are intended to support students' active learning by moving direct instruction out of the classroom to spend in-class time on collaborative work (Cevikbas & Kaiser, 2023; Weinberg et al., 2021). In university proof-based courses, this approach not only applies to the introduction of new mathematical concepts but also to the presentation of proofs (Talbert, 2015). In proof videos, a proof is presented step-by-step by writing down the relevant mathematical steps while the instructor thinks aloud. Doing so, the video gives insight into an expert's behavior (Talbert, 2015) which can help students to better understand mathematical concepts and to learn new methods for proof construction (Mejia-Ramos et al., 2012). However, it is widely known that for many undergraduate students difficulties arise not only in the construction of proofs, but also in reading and understanding the proofs presented (e.g. Roy et al., 2017). In general, presenting mathematics via instructional videos is considered beneficial compared to a face-to-face presentation, because videos enable students to pause and replay and, therefore, allow students to adapt the audio-visual explanations to their pace of learning (Fyfield et al., 2022). However, previous studies indicated that implementing videos into teaching practice is not always beneficial for learning outcomes, but depends on the students' engagement with the videos provided (Chi & Wylie, 2014; Weinberg et al., 2021). Specific research on instructional videos that present proofs is scarce. Yet, a previous study showed that while proof videos were positively perceived by students, their comprehension of the presented proof was quite low (Wirth et al., 2024). To explain these results, in this study, we investigate how students engage in proof videos

and the extent to which their engagement is related to comprehending the presented proof.

THEORETICAL FRAMEWORK

In this section, we briefly introduce the theories at focus, namely proof comprehension (primarily presented in written form) and learning from instructional videos, before relating the two frameworks to each other.

Students' proof comprehension

According to Mejia-Ramos et al. (2012), proper proof comprehension involves both local and holistic understanding of the proof presented. Students demonstrate local comprehension when they understand the meaning of terms and statements, the justification of claims, the logical status of the proof's statements and the proof framework. Holistic comprehension of the proof is indicated when students are able to summarize high-level ideas, identify the modular structure, transfer the general ideas or methods to another context, and understand how the proof relates to examples.

As noted by Hodds et al. (2014), students' proof comprehension depends on how they engage in the proof presented. Thus, students demonstrated enhanced proof comprehension after self-explanation training, where they had to explain each step of a proof, linking it to prior knowledge and previous steps. Similarly, Samkoff and Weber (2015) reported that teaching proof reading strategies, such as 'illustrate the theorem with an example' can support students' proof comprehension. Active engagement with the presented proof appears to be even more important when the written proof is accompanied by additional audio explanations, as is the case with e-proofs. E-proofs supplement a text-based proof on slides with visuals and on-demand audio explanations. A study by Roy et al. (2017) showed that although students felt that e-proofs supported their understanding, they performed worse in a delayed proof comprehension test than those studying a standard written proof. It is likely that students had difficulty integrating the textual and auditory information, because they did not engage in proof comprehension actively (Roy et al., 2017). This assumption is supported by findings on learning from instructional videos.

Students' engagement with instructional videos

The ICAP framework posits that engagement with learning material can influence learning outcomes, categorizing engagement into four modes: *interactive* (e.g., discussing information with a peer), *constructive* (e.g., generating knowledge through self-explanation), *active* (e.g., manipulating learning material by highlighting it) and *passive* (e.g., absorbing information by listening). While passive engagement may result in limited understanding, characterized by verbatim repetition of knowledge, active engagement enables students to apply knowledge on a superficial level. Conversely, constructive or interactive engagement fosters deeper, transferable understanding (Chi & Wylie, 2014). In line with previous research, we assume active or constructive engagement to be reflected in certain observable actions (Weinberg et

al., 2021), namely viewing strategies. Typical viewing strategies are (i) *pausing a video*, (ii) *rewinding certain sections*, and (iii) *taking notes of the video content*.

In a laboratory study, Merkt et al. (2022) demonstrated that learners paused a video due to difficulties in comprehension or the identification of meaningful structural breakpoints within the video. While the detection of gaps in understanding and the rewinding of specific sections to overcome difficulties is regarded as a means to enhance learning, Weinberg et al. (2018) found no association between students' frequency of pausing or skipping back and their learning gains when working with calculus videos. One potential explanation for this finding is that students may encounter difficulties in monitoring their comprehension and, as a result, fail to use the pause and rewind functions effectively. The act of taking notes while watching a video can be indicative of active engagement, particularly when students make verbatim copies of written or spoken text to highlight important aspects. Conversely, taking notes by paraphrasing the text in one's own words can be considered constructive. The latter approach seems more conducive to comprehension, as it involves students connecting aspects of the video to each other or to their own knowledge (Chi & Wylie, 2014). However, empirical research on student note-taking when watching videos is limited. Research on note-taking in mathematical lectures, however, suggests that students often take verbatim notes of what is written on the board (Lew et al., 2016).

RESEARCH QUESTIONS

Based on previous research on both proof comprehension and instructional videos, it is hypothesized that students will demonstrate enhanced proof comprehension test performance if they utilize learning activities associated with an active or constructive mode of engagement. Viewing strategies such as *pausing a video*, *rewinding certain sections*, and *taking notes of the video content* are considered indicators of active and constructive engagement. However, previous research has shown that the frequent use of viewing strategies does not necessarily lead to greater learning gains. The findings on proof comprehension suggest that pausing, rewinding and note-taking may be beneficial when accompanied by self-explanations or proof reading strategies that support active and constructive learning. To our knowledge, there has been no study on strategies for viewing proof videos. Therefore, the present study investigates the extent to which viewing strategies are actually used by students and how viewing strategies can support the understanding of proofs.

RQ1: To what extent do undergraduate mathematics students report using the viewing strategies of pausing a video, rewinding certain sections, and taking notes while watching an instructional video that presents a mathematical proof?

RQ2: To what degree are students' viewing strategies and proof comprehension of the presented proof related?

RQ3: Which reasons for using a viewing strategy do students with low proof comprehension report?

MATERIAL AND METHODS

Data collection took place within a larger project that investigates design criteria for proof videos (Wirth et al., 2024). For this study, we collected data in a real analysis proof-based course for first-semester mathematics (education) students. This course included a traditional lecture and a two-hour tutorial that was conducted as a flipped classroom. In the tutorial, students submitted weekly homework involving proof construction tasks and thereafter received videos in which the proof to be constructed was presented step by step. To achieve meaningful results, we chose to collect data in the middle of the term, when students were already familiar with the procedure.

The selected video presents the proof of the following statement: “Let $f: [0,1] \rightarrow [0,1]$ be a continuous function. Show that f has a fixed point, that is, there exists $x \in [0,1]$ with $f(x) = x$.” Immediately after watching the video, students completed a questionnaire assessing their self-reported viewing strategies. The questionnaire included three single items with a four-point response format. Thus, students rated the extent to which they (i) paused the video to think about it, (ii) rewound certain sections to re-watch them, and (iii) took notes on the video content. To measure students’ proof comprehension, we developed a test with six items based on the assessment model by Mejia-Ramos et al. (2012) with three items targeting the local level and three items targeting the holistic level of proof comprehension. The test had to be completed without accessing the video. A total of 167 students took part in the survey. In addition, a qualitative study was conducted. For this purpose, 19 students were observed on a voluntary basis while watching the video. Based on an observation sheet, a stimulated recall interview was conducted to reconstruct the students’ thoughts and intentions when using viewing strategies (e.g., ‘Why did you stop the video at this point?’).

To answer RQ1 and RQ2, we turned to the descriptive data from the viewing strategies questionnaire and performed correlation analyses to gain insight into the relationship between the self-reported use of viewing strategies and the scores on the proof comprehension test. For RQ3, we selected seven information-rich cases from the interviews for deeper analysis. These cases are characterized by frequent use of viewing strategies but still very low levels of proof comprehension. The qualitative analysis of the students’ interview responses followed the principles of Mayring’s qualitative content analysis (Mayring, 2014). In short, the related themes that emerged were summarized as categories, and these categories were continuously systemized.

RESULTS

RQ1: Table 1 shows the mean values (and standard deviations) of the self-reported use of viewing strategies (in grey). Almost all means roughly correspond to the theoretical scale mean of 2.5, indicating that on average, students stated that they moderately used viewing strategies.

RQ2: The mean scores for proof comprehension (max score: 6) were rather low. Correlation analyses between the self-reported use of viewing strategies and scores on

the proof comprehension test showed no significant ($p < .05$) values (see Table 1). This suggests that the use of viewing strategies does not necessarily lead to better proof comprehension. However, there are highly significant correlations between the individual viewing strategies, indicating that students do not use just one particular strategy, but rather a combination of strategies.

	PC	P	R	TN
Proof comprehension	2.54 (1.56)	-.04	-.11	.05
Pausing		2.81 (1.08)	.48***	.45***
Rewinding			2.39 (1.23)	.25***
Taking notes				2.40 (1.24)

Table 1: Correlations between viewing strategies and proof comprehension with the means (and standard deviations) shown on the diagonal

RQ3: Of the total of 19 students observed, 10 scored only half of the points or less in the proof comprehension test. Of these ten students, three did not use any viewing strategies, which could be due to either a lack of motivation or a lack of assessment of their own understanding. However, the seven others frequently used one or more viewing strategies while watching the video. To provide insight into the data base and the coding, we first discuss the case of Anna, before giving an overview of all cases.

Anna answered two of the six proof comprehension items correctly. These two items captured (1) the main arguments of the proof and (2) the overall proof idea. While watching, Anna paused the video five times and took notes three times. When asked why she paused at the particular points, she mainly refers to the density of information:

Anna (l. 13): “Sometimes a lot of new things appear at once and the person keeps talking. And maybe new things faded in and before the person continues talking, I always pause briefly to understand what's been added, what's new, and yes, I basically read it again, try to understand it, and then, when I think I've understood it, I move on.”

In the situations described by Anna, she appears to monitor her viewing process, recognizes a high cognitive demand, and pauses the video to integrate the visual and auditory information into a coherent understanding. However, the extent to which a deeper processing with self-explanation took place cannot be seen from her comments. In other situations, Anna explained that she uses the pauses to make links to her previous knowledge from the lecture:

Anna (l. 15): I was thinking about which theorem from the lecture says that the difference of continuous functions is continuous again.

In situations like this, Anna also seems to monitor her viewing process and to identify aspects that are not immediately obvious to her, which made her pause the video.

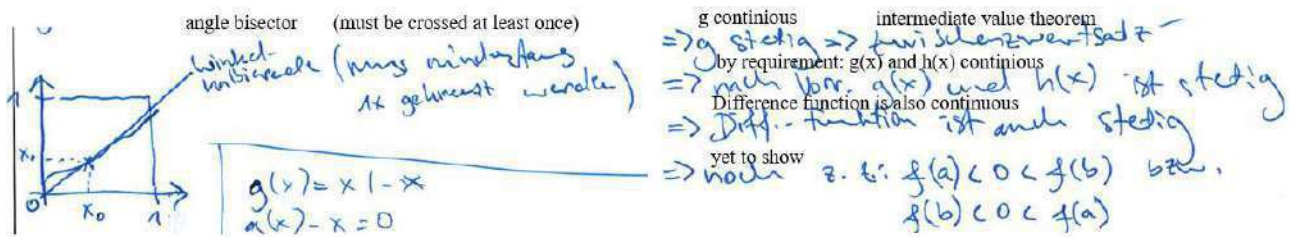


Figure 1: Anna's notes on the proof video

Comparing Anna's notes (Figure 1) to the video, we identified a mix of verbatim quotes from the audio comment and shortened quotes from the visual information. On the left-hand side, she has copied a sketch incompletely but added some explanations from the audio comment in key points. On the right-hand side, she has summarized the argumentation from the video in bullet points (using \Rightarrow as a bullet point and not as an implication). She explained her approach as follows:

Anna (l. 19): Yes, at some point my approach was to first watch a section and then evaluate at the end what was important, what I wanted to take away. And yes, and that's basically what I did, so I read through it again briefly, then looked at what of it is / which individual aspects are important for the chain of argumentation, and then I wrote them down briefly and concisely.

Overall, Anna used the viewing strategies of pausing and note-taking to process information (information density), to question it (link to existing knowledge), and to record it (summarize important aspects). In doing so, she always stuck very closely to the video content and focused on the central arguments of the proof and the general proof idea. It is therefore not surprising that she was able to correctly answer those items from the proof comprehension test that related to these aspects of the proof. She was less likely to question details such as the prerequisites of the intermediate value theorem and was therefore unable to answer the corresponding items correctly.

In analyzing further interviews, we identified a total of nine reasons why students with low proof comprehension used viewing strategies when watching the video. For pausing the video, these are: (1) information density, (2) link to prior knowledge, (3) discovering a gap in understanding, and (4) organizing the video into meaningful sections. When rewinding, all students stated that they (5) could no longer follow the argumentation in the video. When taking notes, the following reasons were given: (6) summarizing important aspects, (7) writing down previously unknown connections, (8) recording information that can be used for other proofs, and (9) learning by writing.

DISCUSSION

In this study, we investigated students' engagement with proof videos and how their viewing strategies relate to proof comprehension. For RQ1, students reported a moderate use of the viewing strategies of pausing the video, rewinding certain sections, and taking notes. As these strategies are associated with active or constructive engagement, according to the ICAP hypothesis (Chi & Wylie, 2014), such viewing strategies could support students' proof comprehension. However, similar to other

studies that found no association between pausing or rewinding and learning gains (Weinberg et al., 2021), this study demonstrated specifically for proof videos that the frequent use of viewing strategies is not sufficient for students to understand the proof presented (RQ2). While students with sound proof comprehension may not have used viewing strategies because they were able to follow the proof presentation in one piece, those cases where students demonstrated poor proof comprehension after frequent use of viewing strategies challenge the ICAP hypothesis. To gain deeper insight into students' viewing strategies, we used qualitative interview data from seven cases and identified nine reasons for students' use of viewing strategies in response to RQ3.

In line with findings by Merkt et al. (2022), students indicated to pause or rewind the video because they experienced difficulty in understanding (due to information density or gaps in understanding) or because they intended to reorganize information (by linking it to prior knowledge or structuring it in meaningful sections). In particular, some students referred to a high cognitive demand resulting from the combination of visual and auditory information, which could pose a similar challenge to working with e-proofs (Roy et al., 2017). In terms of note-taking, students often stuck closely to the content of the video, which is consistent with observations from lectures (Lew et al., 2016). In most cases, they wrote down a mix of verbatim quotes from the audio commentary and shortened quotes from the visual information. In doing so, they paid particular attention to the main arguments of the proof or key ideas.

Overall, in the observed cases, students demonstrated active or constructive engagement with the proof video in terms of the ICAP framework. Moreover, the reasons given for the use of viewing strategies suggest the use of reading strategies such as the elaboration of high-level ideas or the proof method (Samkoff & Weber, 2015). However, these students scored low on the proof comprehension test, because they focused on single components of the proof such as central arguments or key ideas. These components are discussed specifically in the videos, so students may remain at a superficial level of understanding. Thus, trainings covering self-explaining (Hodds et al., 2014) or specific proof reading strategies (Samkoff & Weber, 2015) might be beneficial for these students to move away from the structure of the video and think beyond it. This could also help them to identify gaps in their understanding of implicitly mentioned components of the proof. However, as our results are based on only seven interviews, future studies should investigate students' engagement with proof videos in more detail. Nevertheless, our findings highlight the importance of distinguishing between different levels of active and constructive engagement, ranging from pausing the video due to information density to questioning the details and creating own explanations.

References

- Cevikbas, M., & Kaiser, G. (2023). Can flipped classroom pedagogy offer promising perspectives for mathematics education on pandemic-related issues? A systematic literature review. *ZDM*, 55(1), 177–191. <https://doi.org/10.1007/s11858-022-01388-w>

- Chi, M. T. H., & Wylie, R. (2014). The ICAP Framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, 49(4), 219–243. <https://doi.org/10.1080/00461520.2014.965823>
- Fyfield, M., Henderson, M., & Phillips, M. (2022). Improving instructional video design: A systematic review. *Australasian Journal of Educational Technology*, 150–178. <https://doi.org/10.14742/ajet.7296>
- Hodds, M., Alcock, L., & Inglis, M. (2014). Self-Explanation Training Improves Proof Comprehension. *Journal for Research in Mathematics Education*, 45(1), 62–101.
- Lew, K., Fukawa-Connelly, T. P., Mejia-Ramos, J. P., & Weber, K. (2016). Lectures in advanced mathematics: Why students might not understand what the mathematics professor is trying to convey. *Journal for Research in Mathematics Education*, 47(2), 162–198.
- Mayring, P. (2014). *Qualitative content analysis: Theoretical foundation, basic procedures and software Solution*. <http://nbn-resolving.de/urn:nbn:de:0168-ss0ar-395173>
- Mejia-Ramos, J. P., Fuller, E., Weber, K., Rhoads, K., & Samkoff, A. (2012). An assessment model for proof comprehension in undergraduate mathematics. *Educational Studies in Mathematics*, 79(1), 3–18. <https://doi.org/10.1007/s10649-011-9349-7>
- Merkt, M., Hoppe, A., Bruns, G., Ewerth, R., & Huff, M. (2022). Pushing the button: Why do learners pause online videos? *Computers & Education*, 176, 104355. <https://doi.org/10.1016/j.compedu.2021.104355>
- Roy, S., Inglis, M., & Alcock, L. (2017). Multimedia resources designed to support learning from written proofs: an eye-movement study. *Educational Studies in Mathematics*, 96(2), 249–266. <https://doi.org/10.1007/s10649-017-9754-7>
- Samkoff, A., & Weber, K. (2015). Lessons learned from an instructional intervention on proof comprehension. *The Journal of Mathematical Behavior*, 39, 28–50. <https://doi.org/10.1016/j.jmathb.2015.05.002>
- Talbert, R. (2015). Inverting the transition-to-proof classroom. *PRIMUS*, 25(8), 614–626. <https://doi.org/10.1080/10511970.2015.1050616>
- Weinberg, A., Martin, J., & Tallmann, M. (2021). The effects of instructors and student activity in learning from instructional calculus videos. In D. Olanoff, K. Johnson, & S. Spitzer (Eds.), *Proceedings of the 43rd annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 1733–1737). PMR-NA.
- Weinberg, A., Thomas, M., Martin, J., & Tallmann, M. (2018). Failing to rewind: Students' learning from instructional video. In T. E. Hodges, G. J. Roy, & A. M. Tyminski (Eds.), *Proceedings of the 40th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 1263–1266). PMR-NA.
- Wirth, L., Kirsten, K., Serpé, C., & Greefrath, G. (2024). Comparing different types of instructional videos in a flipped proof-based classroom. *Frontiers in Education*, 9, Article 1438355. <https://doi.org/10.3389/educ.2024.1438355>

THEORIZING OPPORTUNITIES TO LEARN AS A BASIS FOR INCLUSIVE MATHEMATICS EDUCATION

Nancy Kress* and Pablo Duran Oliva+

*University of Colorado Boulder, +Virginia Commonwealth University

In this theoretical essay we propose the adoption of a single definition for Opportunities to Learn (OtL), and we theorize how OtL provides a basis for inclusive mathematics education. We begin by providing our theoretical framework in which we apply a critical perspective that serves to shift attention toward educational and societal systems, structures, resources and power dynamics and away from individuals, their families and communities as the drivers of educational achievement. We describe numerous common ways that OtL has been used, and we propose the adoption of a single robust definition of OtL. We then theorize how this specific definition supports a precise conceptualization of inclusive mathematics education.

INTRODUCTION AND OVERVIEW

Educational achievement gaps between demographic groups is an ongoing focus of attention in undergraduate mathematics (Freeman et al., 2014; Kramer et al., 2023; Theobald et al., 2020). However, in the last few decades the concept of opportunity gap has been resoundingly promoted as a more appropriate focus for improvement efforts in mathematics education (Flores 2007; Carter & Welner, 2013) and researchers in undergraduate mathematics education have increased efforts to understand the impacts of students' learning experiences and opportunities to participate (Johnson et al., 2020; Reinholz et al., 2022; Setren et al., 2021). Focusing on opportunity gaps directs our attention toward educational and societal systems, structures, resources and power dynamics that shape the learning experiences – the opportunities to learn (OtL) – available to students. The focus on unequal OtL decenters individual students with their families and communities from being considered as the source of the problem regarding differences in measured educational achievement. Referring to education broadly, Ladson-Billings, in her 2006 American Educational Research Association presidential address, questioned “the wisdom of focusing on the achievement gap as a way of explaining and understanding the persistent inequality that exists (and has always existed) in our nation’s schools” (Ladson-Billings, 2006, p. 4).

In spite of decades of effort to shift attention toward differences in OtL, paired with efforts to increase and equalize OtL across demographic groups, the concept of OtL itself is incompletely theorized. The term is used to represent constructs ranging from the content included in textbooks (Hadar, 2017) to teachers' professional qualifications (Wang, 1998) to the ways students participate in small groups (Esmonde, 2009a). While each of these constructs is relevant and worthy of consideration in relation to students' OtL, none of them can stand alone as the sole definition of OtL or the driver

of students' academic participation and success. A more robust definition of OtL that contributes to increased understanding of inclusive mathematics education is needed.

THEORETICAL FRAMEWORK

Gutiérrez (2002, 2009) and Martin (2003) called for the adoption of a critical perspective in mathematics education research over twenty years ago. The critical perspective calls for expanding our focus to encompass more than students' academic achievement as measured on traditional tests and directs our attention toward students' access to high-quality resources, classroom learning environments, power dynamics and positionalities. More broadly, the critical perspective calls for us to shift our attention toward societal and systemic factors that impact students' opportunities to learn and have a disproportionately negative effect on the learning experiences of marginalized and underrepresented demographic groups.

Gutiérrez (2009) proposed a definition of equitable mathematics education which directs us to pay attention to constructs represented by two axes: a dominant axis and a critical axis. The dominant axis calls for attending to issues of access and achievement, while the critical axis calls for attending to identity and power dynamics. Attending to equity in mathematics education requires careful consideration of the dominant and critical axes, including access, achievement, identity and power.

In this theoretical essay our application of a critical perspective serves to motivate and justify our adoption of a single definition for OtL. We apply Gutiérrez's (2009) framework for equitable mathematics education to provide explanatory power for our adoption of one specific definition of OtL. Finally, a critical perspective helps to explain how our adopted definition of OtL forms the basis of a precise conceptualization of inclusive mathematics education.

OPPORTUNITIES TO LEARN: A REPRESENTATIVE OVERVIEW OF THE LITERATURE

The concept of opportunity to learn has been applied in different contexts to indicate disparate aspects of students' learning experiences. This section presents representative examples from the literature in which Opportunities to Learn (OtL) are used differently and describes how these varied uses of OtL align with Gutiérrez's dominant and critical axes for equity in mathematics education.

Opportunities to Learn as Access and Achievement

Early conceptualizations of OtL can be found in literature from sixty years ago describing a rudimentary view focused on classroom allocation of time, content exposure, and alignment between teaching and assessment strategies (Anderson, 1967; Carroll, 1963). Moving forward over thirty years, a more comprehensive characterization of OtL adds additional instructional components including content coverage and emphasis and the quality of instructional delivery (Wang, 1998). Even more recently, access to adequate school resources has been recognized as critical to providing meaningful OtL, and research by Bottia et al. (2016) underscores that

students' access to these resources plays a pivotal role in their ability to engage with rigorous academic content and subsequently perform well on traditional measures of achievement.

These conceptualizations of OtL (Anderson, 1967; Carroll, 1963; Wang, 1998; Bottia et al., 2016) are each aligned with the dominant axis of equity in mathematics education as explained by Gutiérrez. They are focused on issues of access to content via resources provided, topics covered and the use of effective teaching methods. In each case, and as supported by Gutiérrez's (2009) framework for equity in mathematics education, these aspects of OtL are shown to be relevant in the role they play in students' learning experiences and opportunities to succeed and achieve in mathematics. They are, however, incomplete as a definition of OtL, since they do not take into consideration how students' identities or the power dynamics at play in mathematics learning communities are impactful in students' experiences of OtL.

Opportunities to Learn as Access to Opportunities to Participate

To capture the most essential aspects of students' experiences inside and outside the classroom, recent conceptualizations of OtL have incorporated student discourse. For instance, Walkowiak (2017) frames OtL as comprising teacher's mathematical knowledge for teaching (MKT), allocation of time during lessons, the mathematical tasks presented to students, and the nature of the mathematical discourse in which students engage. This framework aligns well with Sfard's (2008) Commognition Theory, which emphasizes the importance of discourse and the tools used within it to foster deep learning. Considering learning as an act of participation (Lave & Wenger, 1991), OtL is highly sensitive to students' access to opportunities to participate in the discourse of the mathematical community. This access allows students to engage in academic tasks in a meaningful way, thereby shaping their understanding of mathematics and their ability to use mathematical tools effectively. Access to opportunities to participate in discourse about mathematics is especially relevant to students' experiences as learners in student-centered, collaborative and active learning classrooms.

In a different but related approach, Nachlieli and Tabach (2019) offer another conceptualization of OtL in which they distinguish between "ritual-enabling OtL" that reinforce existing knowledge, and routines and "exploration-requiring OtL" that support and invite students to engage with new, complex, and often unfamiliar content. Both types of OtL are essential, as they provide students with the opportunity to consolidate prior knowledge and explore new concepts.

Incorporation of student discourse and exploration into conceptualizations of OtL aligns with the more interactive, and potentially inquiry-based, nature of student-centered, collaborative and active learning mathematics classrooms. With this shift it becomes increasingly important to consider the role of students' positionalities, identity group memberships, and associated power dynamics in shaping the nature of classroom interactions and students' opportunities to participate. While

conceptualizations of OtL that incorporate student discourse don't often explicitly attend to the critical axis constructs of identity and power, this shift demands that we pay greater attention to how positionality, identity and power dynamics are impactful in students' OtL.

A ROBUST DEFINITION OF OPPORTUNITIES TO LEARN

In this section we present Esmonde's (2009a) conceptualization of OtL, describe how it aligns with the critical axis of Gutiérrez's (2009) framework for equity in mathematics education, and expand on these ideas by sharing Esmonde's (2009b) conceptualization of equitable mathematics education.

Opportunities to Learn in Small Groups

Esmonde (2009a) provides a framework for understanding OtL grounded in sociocultural theory and focused on contexts where students engage in mathematics discourse in cooperative groups. Esmonde's (2009a) framework presents four points, stating that "learning happens (a) through participation, (b) in relation to a social ecology, (c) through processes of identity development, and (d) through communicating about mathematical content" (p. 1011). This definition emphasizes the relational nature of learning environments and how they develop and evolve through classroom interactions and discourse. Esmonde's conceptualization of OtL depends on resources and time being available to students and considers aspects of how those resources and time are taken up and used by instructors and students. Esmonde explicitly attends to the social dynamics of the classroom, emphasizing how access to certain types of learning opportunities is determined by students' interactions with teachers and peers, their positionalities and their associated ability to participate in the mathematical discourse of the community. This view of OtL aligns with a sociocultural perspective on learning (Lerman, 2001) in which OtL is seen as both a function of structural resources (e.g., time, content, instruction) and the participatory mechanisms that allow students to take up the available resources, instruction and time to meaningfully engage with the subject matter.

Summarizing a Robust Definition of Opportunities to Learn

Esmonde's definition of OtL, unlike other common uses of the term, centers identity and power, and is thus also aligned with the axis of Gutiérrez's (2009) framework for equity in mathematics education. As such it provides a definition that aligns with our critical perspective and supports the advancement of our understanding of the nature of inclusive mathematics education which will be explained in detail in the next section of this essay.

However, if we are to take Esmonde's framework for OtL as our robust definition on its own it is crucial that we name the ways that access and achievement come into play in this definition. For students to participate and communicate about mathematical content as described by Esmonde, it is necessary that they have access to resources, time and effective instruction, all of which contribute to OtL and align with the

dominant axis of equitable mathematics education. Thus, Esmonde's framework for OtL depends on the presence of additional forms of OtL as described above under the heading "Opportunities to Learn as Access and Achievement." In essence, Esmonde's version of OtL is complete only if it is assumed that social ecologies encompass students' access to resources, time and effective instruction. Since Esmonde's (2009a) explanation of what is meant by social ecologies is not clearly consistent with this interpretation, we propose naming a fifth component to how learning happens: e) in relation to the availability of resources, time and high-quality instruction.

MOVING BEYOND EQUITABLE: CONCEPTUALIZING INCLUSIVE MATHEMATICS EDUCATION

In this section we reconnect to our critical perspective and affirm our commitment to attending to identity and power in mathematics classrooms. We use Esmonde's definition of OtL and of equitable mathematics education to make explicit the importance of students' participation and ability to move from peripheral to central roles of participation as foundational to inclusive mathematics education. We then interrogate what practices students are participating in when they move into more central roles of participation and suggest dialogue as a mechanism for students' perspectives and ideas to be taken up into the evolving practices of mathematics.

A Definition of Equitable Mathematics Education

Esmonde (2009b) built on the previously described characterization of OtL to define equitable mathematics education as based on the existence of a "fair distribution of opportunities to learn" (p. 249). Esmonde's (2009a) description of the nature of learning is helpful here: Esmonde, citing Lave & Wenger (1991) and Rogoff (2003), stated that the nature of learning is defined by "a change in participation in a set of collective practices (Lave & Wenger, 1991). This change comes about through adaptation to or adoption of a community's ways of speaking, acting, and interacting – although these community practices also change (Rogoff, 2003)" (p. 1011).

Esmonde's (2009b) definition of equitable mathematics education relies on the clear distinction between peripheral and central forms of participation. Peripheral participation depends on peers or the instructor supporting a students' competent participation and/or a student observing practices that are primarily carried out by others. Central participation, on the other hand, refers to participation in which a student contributes to the work of the group and has their ideas and contributions taken up by others in the group. Esmonde (2009a) articulates the importance of students who participate peripherally being able to move into more central positions and not remain in persistently marginal positions, unable to take up more central roles. In summary, Esmonde's definition of equitable mathematics education suggests that fair distribution of opportunities to learn involves forms of student participation that allow movement from peripheral participation toward central participation in classroom practices.

Shifting from Equitable to Inclusive

As we consider students' ability to move into roles of central participation it's important to be clear about what, exactly, students are participating in. Norms of participation, including ways of knowing, doing and using mathematics typically reflect the influences of majority groups. Ong, Smith & Ko (2018) stated that "the prevailing culture and structural manifestations in STEM have traditionally privileged norms of success that favor competitive, individualistic, and solitary practices—norms associated with White male scientists" (p. 206).

Students' movement into more central roles of participation in the existing, established norms of mathematics is reflected in students' "adaptation to or adoption of a community's ways of speaking, acting, and interacting" (Esmonde, 2009a, p. 1011). But, remembering that community practices can, and do, change (Rogoff, 2003), inclusive mathematics education depends on our asking: how do student's ways of thinking about, doing, valuing and using mathematics become incorporated into the changing practices of evolving mathematics communities? Inclusion means that diverse experiences and perspectives are influential in shaping the trajectory of the shifting practices of the community, and over time they become fully and centrally incorporated into the practices of the community

Dialogue as collective sensemaking provides a means for students' ideas and ways of doing mathematics to be taken up in the evolving practices of the community. Dialogue supports learning that is multidirectional. As described by Freire and Macedo (2003) dialogic teaching and learning "characterizes an epistemological relationship" (p. 191), "a process of learning and knowing" (p. 193) in which all participants have agency in the nature of the learning that occurs.

CONCLUSION

Gutiérrez's framework for equitable mathematics education highlights that paying close attention to identity and power in mathematics classrooms may help us learn what helps students who are members of marginalized and/or underrepresented identity groups thrive in mathematics. Esmonde's conceptualization of equitable mathematics education in terms of "fair distribution of opportunities to learn" contributes further to our ability to understand equitable mathematics education. Building off of Gutiérrez's (2009) and Esmonde's (2009b) definitions of equitable mathematics education we theorize inclusive mathematics education in which dialogue provides a mechanism through which students who move into central roles of participation can see their ideas, experiences, and ways of doing and thinking about mathematics taken up and incorporated into the practices of mathematics learning communities.

Acknowledgements

This work was partially supported by the National Science Foundation (NSF) under Grant No. 2400902, No. 2400903 and No. 1934054. Any opinions and conclusions expressed are those of the authors and do not necessarily reflect the views of the NSF.

References

- Anderson, C. A. (1967). The International Comparative Study of Achievement in Mathematics. *Comparative Education Review*, 11(2), 182–196. <https://doi.org/10.1086/445292>
- Bottia, M. C., Valentino, L., Moller, S., Mickelson, R. A., & Stearns, E. (2016). Teacher Collaboration and Latinos/as' Mathematics Achievement Trajectories. *American Journal of Education*, 122(4), 505–535. <https://doi.org/10.1086/687274>
- Carroll, J. B. (1963). A Model of School Learning. *Teachers College Record: The Voice of Scholarship in Education*, 64(8), 1–9. <https://doi.org/10.1177/016146816306400801>
- Carter, P. L., & Welner, K. G. (2013). *Closing the Opportunity Gap: What America Must Do to Give Every Child an Even Chance*. Oxford University Press.
- Esmonde, I. (2009a). Mathematics Learning in Groups: Analyzing Equity in Two Cooperative Activity Structures. *Journal of the Learning Sciences*, 18(2), 247–284. <https://doi.org/10.1080/10508400902797958>
- Esmonde, I. (2009b). Mathematics Learning in Groups: Analyzing Equity in Two Cooperative Activity Structures. *Journal of the Learning Sciences*, 18(2), 247–284. <https://doi.org/10.1080/10508400902797958>
- Flores, A. (2007). Examining disparities in mathematics education: Achievement gap or opportunity gap? *The High School Journal*, 91(1), 29–42.
- Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, 111(23), 8410–8415.
- Gutiérrez, R. (2002). Beyond Essentialism: The Complexity of Language in Teaching Mathematics to Latina/o Students. *American Educational Research Journal*, 39(4), 1047–1088. <https://doi.org/10.3102/000283120390041047>
- Gutiérrez, R. (2009). Framing equity: Helping students “play the game” and “change the game.” *Teaching for Excellence and Equity in Mathematics*, 1(1). <https://journals.charlotte.edu/teem/article/view/1704>
- Hadar, L. L. (2017). Opportunities to learn: Mathematics textbooks and students' achievements. *Studies in Educational Evaluation*, 55, 153–166.
- Johnson, E., Andrews-Larson, C., Keene, K., Melhuish, K., Keller, R., & Fortune, N. (2020). Inquiry and gender inequity in the undergraduate mathematics classroom. *Journal for Research in Mathematics Education*, 51(4), 504–516.
- Kramer, L., Fuller, E., Watson, C., Castillo, A., Oliva, P. D., & Potvin, G. (2023). Establishing a new standard of care for calculus using trials with randomized student allocation. *Science*, 381(6661), 995–998. <https://doi.org/10.1126/science.ade9803>
- Ladson-Billings, G. (2006). From the Achievement Gap to the Education Debt: Understanding Achievement in U.S. Schools. *Educational Researcher*, 35(7), 3–12. <https://doi.org/10.3102/0013189X035007003>

- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.
https://books.google.com/books?hl=en&lr=&id=CAVIOrW3vYAC&oi=fnd&pg=PA11&dq=lave+wenger&ots=ODtEst0EIq&sig=_j2yRFjDy71MF6q97gwbVwEBkyU
- Lerman, S. (2001). Cultural, discursive psychology: A sociocultural approach to studying the teaching and learning of mathematics. *Educational Studies in Mathematics*, 46, 87–113.
- Martin, D. B. (2003). Hidden assumptions and unaddressed questions in mathematics for all rhetoric. *The Mathematics Educator*, 13(2).
<https://openjournals.libs.uga.edu/tme/article/view/1856>
- Nachlieli, T., & Tabach, M. (2019). Ritual-enabling opportunities-to-learn in mathematics classrooms. *Educational Studies in Mathematics*, 101(2), 253–271.
<https://doi.org/10.1007/s10649-018-9848-x>
- Reinholz, D., Johnson, E., Andrews-Larson, C., Stone-Johnstone, A., Smith, J., Mullins, B., Fortune, N., Keene, K., & Shah, N. (2022). When Active Learning Is Inequitable: Women's Participation Predicts Gender Inequities in Mathematical Performance. *Journal for Research in Mathematics Education*, 53(3), 204–226.
<https://doi.org/10.5951/jresmetheduc-2020-0143>
- Rogoff, B. (2003). *The cultural nature of human development*. Oxford, UK: Oxford University Press.
- Setren, E., Greenberg, K., Moore, O., & Yankovich, M. (2021). Effects of flipped classroom instruction: Evidence from a randomized trial. *Education Finance and Policy*, 16(3), 363–387.
- Sfard, A. (2008). *Thinking as communicating: Human development, the growth of discourses, and mathematizing*. Cambridge university press.
[https://books.google.com/books?hl=en&lr=&id=cqwx8aDfUuQC&oi=fnd&pg=PA3&dq=Sfard,+A.+\(2008\).+Thinking+as+communicating:&ots=cBofnTyUoo&sig=JnPb-ikURzqGsFgOAZ9QL1aoR0Y](https://books.google.com/books?hl=en&lr=&id=cqwx8aDfUuQC&oi=fnd&pg=PA3&dq=Sfard,+A.+(2008).+Thinking+as+communicating:&ots=cBofnTyUoo&sig=JnPb-ikURzqGsFgOAZ9QL1aoR0Y)
- Theobald, E. J., Hill, M. J., Tran, E., Agrawal, S., Arroyo, E. N., Behling, S., Chambwe, N., Cintrón, D. L., Cooper, J. D., & Dunster, G. (2020). Active learning narrows achievement gaps for underrepresented students in undergraduate science, technology, engineering, and math. *Proceedings of the National Academy of Sciences*, 117(12), 6476–6483.
- Walkowiak, T. A., Pinter, H. H., & Berry, R. Q. (2017). A reconceptualized framework for 'opportunity to learn' in school mathematics. *Journal of Mathematics Education at Teachers College*, 8(1).
<https://journals-test.library.columbia.edu/index.php/jmetc/article/view/800>
- Wang, J. (1998). Opportunity to Learn: The Impacts and Policy Implications. *Educational Evaluation and Policy Analysis*, 20(3), 137–156.
<https://doi.org/10.3102/01623737020003137>
- Wenger, E. (1998). *Communities of practice: learning, meaning and identity*. Cambridge University Press.

ADHERING TO PRINCIPLES FOR QUALITY TEACHING: MEASURING FACILITATOR ASSESMENTS OF A FICTIONAL PD SITUATION

Malte Lehmann¹, Victoria Shure¹, Sophia Rothering¹, Christin Laschke², Bettina Roesken-Winter¹

¹University of Münster, ²Leibniz Institute for Science and Mathematics Education

Various principles are important for the quality of mathematics teaching, including “student focus and adaptivity” and “enhanced communication”. These are in the focus of the QuaMath facilitator qualification and the teacher professional development program (TPD), respectively. A reliable and economically usable instrument is needed to examine whether facilitators keep focussing on these two principles over the course of the TPD. In this paper, a vignette that captures the two principles is presented and analysed. Using data from $n = 307$ PD facilitators in primary and secondary level, an exploratory factor analysis was carried out to identify the underlying dimensions and reduce the number of items. A confirmatory factor analysis was then applied, indicating that it is empirically possible to distinguish between these two dimensions.

INTRODUCTION

Five principles of high-quality mathematics teaching are fundamental to the *QuaMath* facilitator qualification and teacher professional development program (TPD) (Prediger et al., 2022b): *cognitive demand*, *conceptual focus*, *longitudinal coherence*, *student focus and adaptivity* as well as *enhanced communication*. The principle of *cognitive demand* emphasizes the necessity of engaging students in higher-order thinking processes (Anderson et al., 2021), particularly with respect to knowledge elements relevant for the specific content (Renkl, 2015). *Conceptual focus* draws upon the idea that high-quality mathematics instruction should aim at conceptual understanding, thereby connecting procedures to the underlying concepts (Hierbert & Carpenter, 1992). This ties in with the principle of *longitudinal coherence*, underscoring the importance of organizing content in long-term learning trajectories, connecting new content to prior knowledge, thereby teaching in a manner that fosters longitudinal learning (Prediger et al., 2022b). *Student focus and adaptivity* necessitates that teachers monitor and utilize students’ learning progress, adapting to students’ learning needs (Hardy et al., 2019; Treffers, 1993). Lastly, and in conjunction with the other principles, *enhanced communication* emphasizes the engagement of students in small group or whole-class discussions as crucial for supporting mathematical learning (e.g., Webb et al., 2019).

As the TPD program will be led by facilitators, their ability to effectively incorporate the five principles of high-quality teaching into the PD is decisive. There are limited

findings concerning facilitators' estimation of the principles of *student focus* and *adaptivity* and *enhanced communication* for teachers' learning during PD. The present study aims to fill this gap and present an instrument for measuring facilitators' expertise concerning these two principles. If, after this step, the instrument has proven to be feasible, such an instrument can be developed for the other principles in subsequent studies based on these findings.

THEORETICAL BACKGROUND

Student focus and adaptivity

The *student focus* principle draws from Freudenthal's Realistic Mathematics Education (Freudenthal, 1983), emphasizing that students should engage authentically with mathematics by re-inventing concepts through realistic contexts. Teachers play a key role in supporting this process by closely monitoring students' thinking, using it to inform teaching decisions, and progressively guiding students toward formal mathematical understanding (Treffers, 1993). Teachers mediate between students' ideas and formal mathematics, ensuring active engagement and meaningful learning connections. *Adaptivity* complements *student focus* by addressing individual students' developmental stages and diverse needs. It involves tailoring instruction at both macro (lesson planning) and micro (real-time interactions) levels, ensuring equitable learning opportunities (Hardy et al., 2019). Macro-adaptivity includes differentiating tasks and support structures, while micro-adaptivity focuses on in-the-moment adjustments through questions and prompts.

Enhanced communication

Empirical studies highlighted the importance of mathematical communication, particularly in small-group and whole-class discussions, for fostering students' conceptual understanding and higher-order mathematical practices like reasoning and modelling (Walshaw & Anthony, 2008; Howe et al., 2019). While procedural skills can be learned without communication, rich mathematical discourses enable students to articulate, revise, and refine their thinking, addressing misconceptions and promoting reflective shifts in understanding (Sfard et al., 1998; Ing et al., 2015). Communication is thus essential for maintaining cognitive demand and a conceptual focus in mathematics instruction.

However, research has shown that communicative processes do not inherently support conceptual understanding; students must learn how to communicate effectively. Diverse competencies among students can lead to inequitable participation in discourse practices, emphasizing the need for an *enhanced communication* principle. This principle involves two key aspects: (a) creating rich mathematical discourses to develop concepts and practices and (b) systematically enabling students to participate by promoting their agency, discourse skills, and language abilities (Erath et al., 2021). Teachers play a critical role by fostering students' abilities to justify conclusions, evaluate arguments, and ask clarifying questions. This requires establishing explicit

socio-mathematical norms, scaffolding communication structures, and designing activities with meaningful communicative interdependence (Ing et al., 2015; Webb et al., 2014). *Enhanced Communication* ensures equitable opportunities for all students to engage in and benefit from rich mathematical discourses.

PRESENT STUDY AND RESEARCH QUESTIONS

Facilitators play a decisive role in the enactment of PD program core principles (Borko et al., 2011), and are mainly responsible to maintain the PD content focus over the course of the PD program (Desimone, 2009). In regard to the importance of the two principles *student focus and adaptivity* and *enhanced communication*, it is necessary to develop a reliable and economically usable instrument to measure facilitators' estimation of these two dimensions. On the basis of the previously reported theory, six items were formulated for each of the two dimensions, for which the facilitators were asked to evaluate on a 6-point Likert scale with regard to a given fictional PD situation. Thus, the instrument was aligned with a fictive PD situation to measure facilitators' assessment approximately. The following three research questions were used to test the instrument:

RQ1: Can the facilitators' rating of the items be described by two underlying dimensions that correspond to the scales?

RQ2: Which items contribute most to the assessment of the respective dimension, and can the scales be shortened by removing less informative items without significantly impairing the measurement quality?

RQ3: Can the structure found by the EFA be confirmed by the empirical data of the second continuation?

METHODOLOGY

Participants, context, and data collection

The present study is situated within the large-scale PD program *QuaMath*, which aims to improve PD quality, and, ultimately, mathematics instruction in Germany. The program duration is 10 years, with an outreach of approximately 10,000 schools from primary to secondary level. The focus on instructional quality is guided by five principles for mathematics teaching, gained by an extensive literature review: *conceptual focus*, *cognitive demand*, *student focus and adaptivity*, *longitudinal coherence*, and *enhanced communication* (Prediger et al., 2022). First, facilitators attend a one-year PD program dedicated to exploring the five principles in detail for different mathematics topics, and how to support teacher learning in this respect. Second, they then provide the PD courses for teachers themselves. Particularly, the facilitator PD aims to strengthen their expertise in relation to teachers' learning of the five principles.

PD facilitators from 15 out of 16 Federal States in Germany started the program in September 2023. Among them, $n = 139$ facilitators from the primary level and $n = 168$

from the secondary level, with $M = 8.9$ ($SD = 6.8$) years of mathematics teaching experience, opted to participate in this study. Facilitation experience ranged anywhere from 0 to more than 7 PD sessions held as facilitators. Prior to the start of the program, the facilitators completed a survey containing demographic questions as well as a situated assessment of their expertise in regard to teacher utterances concerning the main principles of the program. Thus, respective aspects were embedded in a fictional dialogue of teachers from a PD, discussing the use of a digital learning application (app) in the classroom (Figure 1). In this fictional dialogue, the teachers primarily discuss short-term and motivational benefits of the use of the learning app.

Situation in a mathematics PD:

Ms. Miller is raving about a new app that she used in her class:

Ms. Miller: The app is fantastic. My students can work exactly on their level, because the multiplication tasks will always be adjusted.

Ms. Özdemir: I know that app too. Then they just sit there in front of their tablets and practice multiplication. Although it is nice and quiet.

Mr. Meier: I also use it, everyone can work on his or her own level and just get the tasks that they need. And my students are really proud of themselves when they figure out lots of tasks and they have more confidence for the class test.

Figure 1: PD situation - dialogue amongst teachers concerning the use of an app.

The facilitators were asked to assess two possible continuations of the PD situation (see Figure 2, for one of the continuations) with regard to the two principles of *student focus and adaptivity* and *enhanced communication*.

Continuation 2 was composed to reflect a facilitator aiming at supporting teachers' learning of the two principles, whereas continuation 1 did not reflect this. For both continuations, the participants were asked to assess, on the basis of six items, the extent to which the facilitator had taken the two quality principles into account. Two example items are "The facilitator elicits that the teachers talk about the potential of the app for a joint discussion" (*enhanced communication*) and "The facilitator elicits that the teachers discuss the app's potential for learning support" (*student focus and adaptivity*). All items have been assessed on a 6-point Likert scale, ranging from "disagree completely" to "agree completely".

<p>Continuation 1 of the PD situation</p> <p>Facilitator: You all addressed an important aspect, namely that the students can control their own learning with the app and focus on the things they need. What else is important to you when considering using the app?</p> <p>Mr. Meier: Above all, the experience of success is important.</p> <p>Facilitator: That's another nice thing about an app like this.</p> <p>Ms. Özdemir: Well, you can do that without an app, that's how I prefer to do it.</p> <p>Ms. Müller: However, the app asks questions in a dynamic way, depending on the individual learning level. This makes my work much easier.</p> <p>Facilitator: Yes, it's not an easy call. Ms. Özdemir, if you don't have the app, how do you deal with it?</p> <p>Ms. Özdemir: My students are proud when they can successfully solve various multiplication problems, for instance. But they don't have to work with an app to do this. I select worksheets that are individually appropriate.</p> <p>Facilitator: How do the others see it?</p> <p>Mr. Meier: However, selecting individual worksheets is quite time-consuming - so I think you save a lot of time with the app.</p> <p>Facilitator: That's a good point.</p>

Figure 2: Continuations 1 of the PD situation - dialogue amongst facilitator and teachers concerning the use of an app.

Data analysis

For each item, the 6-point Likert scale was converted to 0 to 5 points. For RQ1 and RQ2, an exploratory factor analysis (EFA) was performed on facilitators' responses to continuation 1 to determine the factors contributing to the principles that can enhance PD quality using SPSS (version 29.0). Both a Bartlett test ($\chi^2(66) = 1160.65, p < .001$) and a Kaiser-Meyer-Olkin test ($KMO = .851$) indicated the variables to be suitable for factor analysis. A principal component analysis with varimax rotation was therefore carried out. Eigenvalues were used to determine the number of factors (those with values exceeding 1). The items with factor loadings' absolute values less than 0.4 and items with two similar factor loadings were deleted, and the remaining items were used to rerun EFA. Finally, a reliability analysis was carried out for both shortened scales.

For RQ3, a confirmatory factor analysis (CFA) was performed on facilitators' responses to continuation 2 to test the model found in the EFA using the R-package lavaan (Rosseel, 2012).

RESULTS

In regard to RQ1 and RQ2, the final run of EFA was performed on six items regarding the two principles resulting in two factors that explain a total variance of 61.1%. The first factor, *facilitators' assessment of student focus and adaptivity*, comprised four items with factor loadings ranging from 0.670 to 0.845. The internal consistency of the scale is satisfying with Cronbach's $\alpha = .77$. The second factor, *facilitators'*

assessment of enhanced communication, comprised four items with factor loadings ranging from 0.700 to 0.838. The internal consistency of the scale is also satisfactory, indicated by Cronbach's $\alpha = .76$.

	Factor loading	
	1	2
Factor 1: Facilitators' assessment of student focus and adaptivity		
The facilitator elicits that the teachers discuss the app's potential for learning support.	.845	.177
The facilitator elicits that the teachers talk about the suitability of the app for supporting learning processes.	.812	.211
The facilitator elicits that the teachers address the possibilities of the app in supporting learning.	.682	.006
The facilitator elicits that the teachers talk about the benefits of the app for learning development.	.670	.265
Factor 2: Facilitators' assessment of enhanced communication		
The facilitator elicits that the teachers discuss the possibility of conversations when using the app.	.163	.838
The facilitator elicits that the teachers talk about the app's suitability for communicating with each other.	.119	.781
The facilitator elicits that the teachers talk about the potential of the app for a joint discussion.	.124	.747
The facilitator elicits that the teachers address the suitability of the app for starting conversations.	.191	.700

Table 1: Items and their loadings for the two factors.

For RQ3, CFA was conducted to examine the two factors. Based on the results of the EFA, we hypothesized two latent factors: *facilitators' assessment of enhanced communication* (EC) and *facilitators' assessment of student focus and adaptivity* (SA). The model fit indices suggested a good fit. The Chi-Square value for the model was not significant, $\chi^2(16) = 20.36$, $p = .21$, with a CMIN/DF ratio of 1.27, indicating a reasonable fit relative to the degrees of freedom. The Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI) were both .99, both revealing a strong model fit. Additionally, the Root Mean Square Error of Approximation (RSMEA) was .03 with

a 90% confidence interval of 0.000 to 0.069, so that the model fits well in terms of error approximation.

DISCUSSION AND CONCLUSION

The results of the exploratory factor analysis showed that it is possible to empirically differentiate the facilitators' ratings by underlying dimensions that correspond to the scales, representing two principles for quality teaching. With the resulting four items per dimension, it is possible to measure these reliably and economically. The results of the confirmatory factor analysis ultimately confirmed this structure. The fit indices emphasize that the model fits well with the empirical data. Finally, we were able to develop of a reliable and economically useable instrument for measuring facilitators' assessments of these two principles approximately, but in a situated manner. We will also implement the instrument at the end of the PD qualification program to reveal facilitators' development over time. Not all principles were included in the analysis; for cognitive demand, conceptual focus, and longitudinal coherence, further studies should follow. However, an instrument for attending to student focus and adaptivity, as well as enhanced communication, is a first step towards closing the gap that has existed so far. Instruments for the other principles can be developed accordingly.

Additional information

This research was funded within the project *QuaMath*, financed by the Standing Conference of the Ministers of Education and Cultural Affairs (KMK) in Germany.

References

- Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., et al. (eds) (2001). *A taxonomy for learning, teaching, and assessing: A revision of bloom's taxonomy of educational objectives*. Longman.
- Borko, H., Koellner, K., Jacobs, J., & Seago, N. (2011). Using video representations of teaching in practice-based professional development programs. *ZDM-Math Education*, 43(1), 175–187. <https://doi.org/10.1007/s11858-010-0302-5>
- Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, 38(3), 181–200. <https://doi.org/10.3102/0013189X08331140>.
- Erath, K., Ingram, J., Moschkovich, J., & Prediger, S. (2021). Designing and enacting instruction that enhances language for mathematics learning – A review of the state of development and research. *ZDM-Math Education*, 53, 345–262. <https://doi.org/10.1007/s11858-020-0123-2>
- Freudenthal, H. (1983). *Didactical phenomenology of mathematical structures*. Netherland:Dordrecht.
- Hardy, I., Decristan, J., and Klieme, E. (2019). Adaptive teaching in research on learning and instruction. *J. Educ. Res. Online* 11, 169–191.

- Hiebert, J., and Carpenter, T. P. (1992). "Learning and teaching with understanding," in *Handbook of research on mathematics teaching and learning*, ed. D. A. Grouws (London: Macmillan), 65–97.
- Howe, C., Hennessy, S., Mercer, N., Vrikki, M., & Wheatley, L. (2019). Teacher-student dialogue during classroom teaching: Does it really impact on student outcomes? *Journal of the Learning Sciences*, 28(4-5), 462–512. <https://doi.org/10.1080/10508406.2019.1573730>
- Ing, M., Webb, N., Franke, M., Turrou, A., Wong, J., Shin, N., & Fernandez, C. (2015). Student participation in elementary mathematics classrooms: The missing link between teacher practices and student achievement? *Educational Studies in Mathematics*, 90, 341–356. <https://doi.org/10.1007/s10649-015-9625-z>
- Prediger, S., Roesken-Winter, B., Stahnke, R., & Pöhler, B. (2022). Conceptualizing content-related PD facilitator expertise. *Journal of Mathematics Teacher Education*, 25, 403–428. <https://doi.org/10.1007/s10857-021-09497-1>
- Prediger, S., Götze, D., Holzäpfel, L., Rösken-Winter, B., & Selter, C. (2022). Five principles for high-quality mathematics teaching. *Frontiers in Education*, 7(969212), 1–15. <https://doi.org/10.3389/feduc.2022.969212>.
- Renkl, A. (2015). Different roads lead to Rome: The case of principle-based cognitive skills. *Learn. Res. Pract.* 1, 79–90. <https://doi.org/10.1080/23735082.2015.994255>
- Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48(2), 1–36. <http://www.jstatsoft.org/v48/i02/>
- Sfard, A., Nesher, P., Streefland, L., Cobb, P., and Mason, J. (1998). Learning mathematics through conversation: Is it as good as they say? *For the Learning of Mathematics*, 18(1), 41–51.
- Treffers, A. (1993). Wiskobas and Freudenthal: Realistic mathematics education. *Educ. Stud. Math.* 25, 89–108. <https://doi.org/10.1007/BF01274104>
- Walshaw, M., and Anthony, G. (2008). The teacher's role in classroom discourse: A review of recent research into mathematics classrooms. *Rev. Educ. Res.* 78, 516–551. doi: 10.3102/0034654308320292
- Webb, N., Franke, M., Ing, M., Wong, J., Fernandez, C., Shin, N., & Turrou, A. (2014). Engaging with others mathematical ideas: Interrelationships among student participation, teachers' instructional practices, and learning, 63, 79–93. <https://doi.org/10.1016/j.ijer.2013.02.001>
- Webb, N. M., Franke, M. L., Ing, M., Turrou, A. C., Johnson, N. C., and Zimmerman, J. (2019). Teacher practices that promote productive dialogue and learning in mathematics classrooms. *Int. J. Educ. Res.* 97, 176–186. <https://doi.org/10.1016/j.ijer.2017.07.009>
- Zaslavsky, O. & Leikin, R. (1999). Interweaving the training of mathematics teacher educators and the professional development of mathematics teachers. In O. Zaslavsky (Ed.), *Proceedings of the 23rd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 143–158). PME.

‘MATHEMATICAL TALK VIDEOS’: EMPIRICALLY INITIATED CONCEPTUAL CONSIDERATIONS

Lena Leoci and Rose F. Vogel

Goethe University, Frankfurt am Main, Germany

This paper focuses on the development of central features in the design of mathematical talk videos. Based on the target settings and the context of application for this type of video, theory and empirically based features for the design of such videos are developed and formulated.

INTRODUCTION

This article focuses on *mathematical talk videos*, which are being developed and investigated in the MaLeLiOS (Mathematical Learning Situations – Listening, Observing and Speaking) project for the design of mathematical teaching and learning processes (Scheibelein & Vogel, in press). The aim is to develop a type of video that offers the opportunity to accompany the intertwined differentiation of mathematical and linguistic skills in children.

Based on the mathematics-education assumption that mathematical teaching and learning processes take place in dialogical exchange with other learners (Krummheuer, 2000), a video structure is to be developed that does not explain mathematical facts like an explanatory video but rather can be used as an initiator and orientation for conducting mathematical discussion in elementary-school mathematics lessons. In discursive debate with others, different perspectives on mathematical issues can be adopted. These negotiation processes stimulate and accompany individual mathematical cognitive processes. From a conversational-linguistics perspective, the dialogue structure of narration is to be focused on this video structure (Quasthoff, 2001).

The videos used and analysed for these conceptual considerations were produced in a seminar as part of primary-school teaching-degree program at Goethe University Frankfurt. The students learned the characteristic features of a talk video derived from the target perspective from the seminar leader to incorporate these into the videos they produced. In addition, they received a mathematics-education introduction to the topics of *mathematical talk*, *linguistic narrative structures* and *diagrammatical work*, which describes materialised mathematical work from a semiotic perspective. The specified mathematical topics were assigned to the content areas of (1) patterns and structures, (2) numbers and operations, (3) space and form and (4) data and probability. After the students created an initial version of a script, this was further differentiated and optimised based on feedback from the seminar group and its leader. In the end, the students produced the videos themselves. The aim of analysing the nine *mathematical*

talk videos created by the pre-service teachers, is to empirically work out how the characteristics of a *mathematical talk video* derived from the objective can be staged. The videos created and analysed in dialogue with the pre-service teachers can be used as a preliminary study for the MaLeLiOS project. In the main study, *mathematical talk videos* are compared to illustrated stories in learning environments with third and fourth graders in a longitudinal study.

THEORETICAL BACKGROUND

At the centre of the *mathematical talk video* format is conversation. The mathematical talks staged in the video are intended to have a narrative character on the one hand and to show discursive elements in the dialogical discussion on the other.

The discursive dimension addresses, amongst other aspects, interactional patterns and routines, and in particular discourse practices such as explaining, justifying, arguing, etc. (Erath et al., 2021, p. 246)

The focus on learner participation in mathematical discourses supports mathematical learning, as mathematical negotiation processes take place among the participants, so individual mathematical knowledge can be restructured and enriched (Henschen et al., 2022). Learners can be introduced to discursive practices and participate in processes of knowledge constitution (Erath et al., 2018).

The design of the mathematical teaching–learning situation into which the *mathematical talk video* is integrated stipulates that the discussion conducted in the video is continued by the children after watching it. This means that the video must end with an open-ended situation that needs further clarification and that can be differentiated with linguistics and, if necessary, the materials, or rather artefacts (materials in the sense that they access proven practices within mathematical learning processes: e.g., a place-value table) introduced in the video and led to a solution by the children. The mathematical discussions among the learners can then lead to a learning sequence in which the mathematical knowledge is recorded.

If one assumes that mathematical learning is a social process, interaction processes come into focus. Mathematical content is negotiated in interaction with others, an interaction that is usually narrative in nature for young learners (Krummheuer, 2000).

This interaction is of an argumentative nature and argumentation dissolves in a narrative presentation. This means that: the narrative classroom culture of primary education is based on rationality, and the social constitution of classroom learning is participation in the interactional accomplishment of argumentative, narratively-structured sequences of actions. (Krummheuer, 2000, p. 24)

Children tell a story about their mathematical activities. This story is presented to other children and then develops further or is integrated into another story. The stories often refer to actions that have happened in a situation, and an attempt is made to reproduce them as accurately as possible (Krummheuer, 2000). These stories are usually told in

everyday language and, depending on the level of knowledge and dialog partner, are enriched with mathematical terms.

From a linguistic perspective, narration can be described as a dynamic process that is staged and continued by the narrators and listeners (Quasthoff, 2001). For this to succeed, Quasthoff (2001) formulates tasks that are taken on by the actors involved. The “narrative discourse unit” (Quasthoff, 2001, p. 1300) is opened with the task of “thematizing” (Quasthoff, 2001, p. 1302). This means that the previous interaction process (“turn-by-turn-talk”; Quasthoff, 2001, p. 1302) is abandoned. The task of “elaborating or dramatizing” (Quasthoff, 2001, p. 1302) unfolds the events. Here, a deepening of the topic takes place on both the linguistic and the action or inscriptional levels. In the dramaturgy of the talk video, this means that technical language can be offered at this point, as well as the handling of materials or artefacts, and that inscriptions can be introduced or possible mathematical relations in the presented diagrams can be explained. The “conclusion” ends the “discourse world of history” (Quasthoff, 2001, p. 1303). The beginning (“representation of content relevance”; Quasthoff, 2001, p. 1302) and the end of a discourse unit (“transition”; Quasthoff, 2001, p. 1302) have the following functions:

In one case, it is semantically about the thematic preparation and the representation of content-related connectivity of the discourse world of history [...] and structurally about leaving the turn-by-turn-talk. In the other case, it is about leaving the reference points of the story and structurally about regaining the turn-by-turn-talk. (Quasthoff, 2001, p. 1303, translated by the authors)

The narrative arrangement in the *mathematical talk videos* and their continuation after watching the video opens the possibility of initially indirectly addressing and negotiating mathematical relationships in everyday and fictitious situations (Vogler, 2019) that can then be made increasingly explicit throughout the learning environment.

Mathematical learning is often accompanied using materials, or rather artefacts. Materials and artefacts have various functions. They enable material-based mathematical relationships to be represented and thus offer starting points for talking about mathematics through their material representation. At the same time, the selected object constellations can become the starting point for changes that manifest themselves in manipulations of the object constellation. In a semiotic sense, doing mathematics becomes a sign activity. In this context, Dörfler (2007) formulates that:

One important aspect of mathematics as a sign activity is that it produces symbolic structures that can be used to model situations and processes of many sorts. (p. 106)

This type of sign activity aims “to work out a certain form of materiality and perceptual objectivity in mathematical activities, which the American philosopher Ch.S. Peirce called diagrammaticity.” (Dörfler, 2006, p. 201, translated by the authors).

Diagrams are thus material or inscriptional structures between whose elements can be described rules, which in turn provide the basis for the “manipulations”,

“experimentation”, “analysis of relation” and “invent and design” of new diagrams and “use of diagrams” for modelling (Dörfler, 2006, pp. 213–214).

Diagrams are always organized in systems and a diagram system results in a kind of calculation [...], which includes a practice of operating with the diagrams. (Dörfler, 2014, p. 5, translated by the authors)

Mathematical talk videos can become crystallisation points for mathematical conversations among children if various features are implemented in their design. The theoretical considerations described above provide initial indications for their design and lead to three characteristics (Scheibelein & Vogel, in press), of which four have been described so far:

(1) Focusing on storytelling opens the space for a narrative and therefore initially informal approach to mathematical content. Mathematical phenomena are integrated into everyday situations, fictional situations or mathematical narrative situations, providing a low-threshold introduction to talking about mathematical issues.

(2) The conversations or narrative monologues shown in the *mathematical talk videos* lend the videos the character of a language model. Listening gives learners the opportunity to participate in the discussion that follows the videos, regardless of their language and discourse skills. Words and sentences that are important in describing the mathematical content can be included by the children in their conversations following the video. If different language registers, such as everyday language, educational language and technical language, are used in the *mathematical talk videos*, elements of them can be adopted by the children.

(3) *Mathematical talk videos* provide the opportunity to show how mathematical materials and artefacts can be used. By observing their handling, children can use them for further mathematical considerations. At the same time, they can be shown how to talk about their use. These materials or artefacts can serve as a link between the video and the subsequent conversation among the learners.

ANALYSIS WITH SELECTED RESULTS

The aim of the analysis is to gain empirically based indications for the detailed formulation of the characteristic features of a *mathematical talk video*.

For this purpose, the *mathematical talk videos* created by pre-service education students were evaluated using qualitative content analysis according to Mayring (2014). The main category and subcategories are developed deductively based on theoretical considerations and inductively expanded through the work on the data material. Six main categories can be identified in this way: (1) characterisation of the situation, (2) dialogue form, (3) mathematical arrangement in the sense of the explicitness of mathematical statements, (4) language action (5) actions with objects or gestures and (6) image (iconic representation), which will be further differentiated into subcategories.

For the analysis, each subcategory is described in detail and illustrated with an anchor example from the data material. In total, nine *mathematical talk videos* are coded using this category system. The videos are divided into 10-second coding units, each of which was assigned to a subcategory of the six main categories. In this way, a timeline is created for each video that shows changes or rather stable specifications of the six main categories throughout the duration of the video.

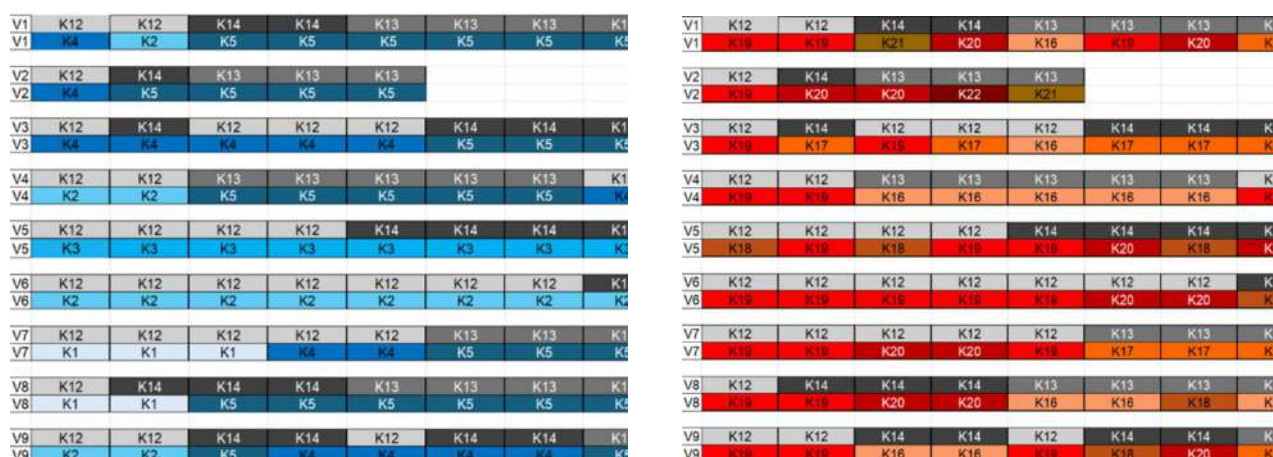
As the central design elements cannot be seen independently of each other, selected main categories are presented in relation to one another here.

Connection: characterisation of the situation – mathematical arrangement in the sense of the explicitness of mathematical statements (see Fig. 1, left)

Characterisation of the situation: everyday situations, such as leisure time or eating (K1–K3), artificial situations (K4) or mathematical situations (K5).

Mathematical arrangement in the sense of the explicitness of mathematical statements: none (K12), explicit (K13) or implicit (K14).

Fig. 1: Excerpts from the comparison of selected main categories at the subcategory level.



On the left side: characterisation of the situation compared to the mathematical arrangement. On the right side: the mathematical arrangement compared to speech.

At the beginning of the videos, everyday situations (K1–K3) dominate. An explicit mathematical arrangement is not identifiable. In some cases, it will be given a hint of mathematical relations, so an implicit reference of a mathematical arrangement is available. In most cases, there quickly occurs a change to an explicit mathematical situation (K5). In summary, it is noticeable that, on the one hand, an explicit mathematical arrangement (K13) is preceded by an implicit mathematical one (K14). There are also videos in which a non-mathematical situation leads directly to an explicit mathematical situation. The implicit design of the situation is skipped.

Overall, the required mathematical arrangements are implemented in the described manner in seven of the nine videos. No explicit mathematical modelling is carried out for the space and form content area because the materials from this everyday situation

(spatially optimal packaging for lunch boxes) can themselves be used for mathematical manipulations.

Implications for the design of the mathematical talk videos:

The narrative situation in the mathematics videos is initiated by a situation from the children's everyday life or imaginary worlds. In this way, the children are targeted in their everyday thought patterns and motivated for the situation. At the same time, the situation unfolds linguistically and with the help of moving images, thereby creating sequential language models in everyday language. The implicit mathematical content is then modelled explicitly with the help of materials, or rather artefacts, so that these signs can be manipulated in the sense of a diagrammatical work and argumentation to obtain mathematical findings. Exemplary manipulations are accompanied by speech and actions with material.

Connection: mathematical arrangement in the sense of the explicitness of mathematical statements – speech (see Fig. 1, right)

Speech: not present (K15), explanations (K16), definitions (K17), confirm understanding (K18), descriptions (K19), exploring (K20), formulating work order (K21), summaries (K22).

Overall, the videos designed by pre-service education students show that describing (K19) as a language action is used more at the beginning of the videos and in phases in which no implicit or explicit mathematical interpretation is yet available. Defining (K17) as a language action and exploring (K20) are often used by pre-service education students in connection with an implicit mathematical interpretation. In these cases, the mathematical interpretation arises from the context of the described situation, and explicit modelling is not considered as important. When explicit modelling is used, explanations (K16) or definitions (K17) are often in the foreground. Overall, seven of nine pre-service education student groups ended their videos with a work order (K21).

Implications for the design of the mathematical talk videos:

The form of mathematical modelling (none, implicit or explicit) should be considered in the arrangement of language actions. Central language actions important to the mathematical process of exploration, such as describing, exploring, defining and explaining, should be considered in the design of the *mathematical talk videos*.

Connection: mathematical arrangement in the sense of the explicitness of mathematical statements – dialogue forms

The narrative introduction is often presented by the pre-service education students in form of a monologue: a story is being told. The focus here is on the context (everyday life, fiction or a mathematical situation). The connections within the context are established and function as the basis for implicit and explicit mathematical interpretations. The chronological sequence following the explicit mathematical interpretation is usually treated in the students' video concepts as a mathematical dialogue in which the interlocutors have equally important speaking parts and the

mathematical modelling is further differentiated during their turns. According to Quasthoff (2001, p. 1303), the task of “elaborating and dramatizing” can be distributed equally among the actors, or one person can be assigned a more dominant role in the talk. This could clarify the expert status of a person, or individual aspects could be represented alternately by the actors involved. In this way, discursive approaches can be worked out exemplarily.

Connection: action – speech

In most cases, the pre-service education students did not present actions but rather accompanied their images with monologues and dialogues. When actions are presented, the focus is on diagrammatical work. The manipulations of object arrangements and the experimentation are mainly presented in a results-oriented way and linguistically accompanied by definitions and explanations. Here the use of rules for the manipulations are the focus.

CONCLUSION AND OUTLOOK

The students’ first attempts at designing their videos show approaches to implementing the three areas characteristic of the *mathematical talk videos*. All student groups chose an everyday, fictitious or mathematical situation. Here, the mathematical area proves to be a central factor that determines the context of the story being told. For a planned *mathematical talk video* for the main study of project MaLeLiOS from the mathematical content area of graph theory, for example, an everyday situation lends itself to the search for possible variants of passability (Euler- and Hamilton-ways). The aim is to find different bus routes in path networks that must be found for certain requirements. Here, explicit mathematical modelling, like the pre-service teacher video about space and form (the optimal packaging of food for lunch boxes) will not be necessary. Furthermore, the analyses show that, in contrast to monologues, a dialogical narrative structure is better suited to exemplify in the *mathematical talk video* how mathematical modelling is negotiated and how the mathematical content is deepened in dialogue.

The trials show that fascinating mathematical tasks, such as recognising and determining the number of dots in different arrangements without counting (the patterns and structures content area), encourage children to solve the task independently in dialogue. The available material is well chosen in this example, as it provides structure on the one hand, while on the other hand the task is still sufficiently challenging. It also became clear that the problem in the video was not completely solved, so the continuation of the discussion about the solving process was attractive for the children. The transition between the video and the talks of the children after observing the video is certainly the biggest challenge in the production of *mathematical talk videos*.

References

- Dörfler, W. (2006). Diagramme und Mathematikunterricht [Diagrams and mathematics education]. *Journal für Mathematik-Didaktik*, 27(3/4), 200–219.
- Dörfler, W. (2007). Making mathematics more mundane – A semiotic approach. In J.-H. Woo, H.-Ch. Lew, K.-S. P & D.-Y. Seo (Eds.), *Proc. 31th Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 1, pp. 105–108). PME.
- Dörfler, W. (2014). 1. Abstrakte Mathematik und Computer [Abstract Mathematics and Computers]. In T. Wassong, D. Frischmeier, P. R. Fischer, R. Hochmuth, & P. Bender (Eds.), *Mit Werkzeugen der Mathematik und Stochastik lernen – Using Tools for Learning Mathematics and Statistics* (pp. 1–14). Springer Spektrum. <https://doi.org/10.1007/978-3-658-03104-6>
- Erath, K., Ingram, J., Moschkovich, J., & Prediger, S. (2021). Designing and enacting instruction that enhances language for mathematics learning: A review of the state of development and research. *ZDM – Mathematics Education*, 53, 245–262.
- Erath, K., Prediger, S., Quasthoff, U., & Heller, V. (2018). Discourse competence as important part of academic language proficiency in mathematics classrooms: The case of explaining to learn and learning to explain. *Educational Studies in Mathematics*, 99(2), 161–179.
- Henschen, E., Teschner, M., & Vogler, A.-M. (2022). Peer interactions and their role in early mathematical learning in kindergarten discourses. *International Electronic Journal of Mathematics Education*, 17(4), em0709. <https://doi.org/10.29333/iejme/12362>
- Krummheuer, G. (2000). Mathematics learning in narrative classroom cultures: Studies of argumentation of primary mathematics education. *For the Learning of Mathematics*, 20(1), 22–32.
- Mayring, P. (2014). *Qualitative content analysis: Theoretical foundation, basic procedures and software solution*. Beltz. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-395173>
- Quasthoff, U. M. (2001). Erzählen als interaktive Gesprächsstruktur [Narration as an Interactive Conversation Structure]. In K. Brinker, G. Antos, W. Heinemann, & S. F. Sager (Eds.), *Text- und Gesprächslinguistik. Ein internationales Handbuch zeitgenössischer Forschung* (2. Halbband) (pp. 1293–1309). De Gruyter Mouton. <https://doi.org/10.1515/9783110169188.2.18.1293>
- Scheiblein, L., & Vogel, R. (in press). Mathematical learning videos as an initiation of a conversation. *Proceedings of the 15th International Congress on Mathematical Education (ICME)*, Sydney, Australia.
- Vogler, A.-M. (2019). Chances and obstacles of ‘indirect’ learning processes in situations with preschool teachers. *Eleventh Congress of the European Society for Research in Mathematics Education (CERME11)*, Utrecht University, Utrecht, Netherlands. HAL Id: hal-02414992.

HOW DO PARENTAL AND TEACHER SUPPORT INFLUENCE HIGH SCHOOL STUDENTS' MATHEMATICS LEARNING MOTIVATION? — A CHAIN MEDIATION MODEL BASED ON GROWTH MINDSET AND SELF-EFFICACY

Xinxin Li¹², Tommy Tanu Wijaya¹², Yiming Cao¹²

¹School of Mathematical Sciences, Beijing Normal University, 100875, China

²National Research Institute for Mathematics Teaching Materials, 100875, China

Mathematics learning motivation significantly impacts students' academic achievement. After the National College Entrance Examination, students' learning motivation relies more on intrinsic psychological traits rather than external pressures. Based on 336 samples, this study uses Partial Least Squares Structural Equation Modeling to examine how external support, and internal factors interact to enhance mathematics learning motivation. Results show: (1) students exhibit relatively high mathematics learning motivation; (2) parental and teacher support indirectly influence motivation via internal mechanisms; (3) growth mindset and self-efficacy, including their chained effect, mediate these relationships, acting as crucial bridges. Home-school collaboration is recommended to boost growth mindset, self-efficacy.

1 INTRODUCTION

Learning motivation is a critical factor influencing students' academic achievement and learning behaviors (Suharnadi et al., 2024). It is commonly defined as the driving force behind sustained engagement in learning activities (Deci & Ryan, 2000). Intrinsic motivation reflects interest and internal satisfaction, while extrinsic motivation relates to external pressures (Suharnadi et al., 2024). Mathematics, as a foundational discipline, significantly influences students' academic development and learning outcomes. Mathematics learning motivation has been shown to significantly impact students' mathematical performance (El-Adl & Alkharusi, 2020). PISA 2022 highlights the importance of enhancing intrinsic and instrumental motivation to cultivate proactive, independent lifelong learners (OECD, 2024).

After the National College Entrance Examination (NCEE), students' mathematics learning motivation shifts from being externally driven to relying more on intrinsic interest and long-term goals, reflecting foundational education outcomes. High school graduates, positioned at the intersection of basic and higher education, represent a key cohort whose motivation shapes their academic and career trajectories (Aeschlimann et al., 2016). Despite extensive research on learning motivation, studies focusing on this transitional group remain limited. Understanding their motivation mechanisms offers valuable insights into fostering learning motivation in both foundational and higher education.

Self-determination theory highlights the significant role of external support, particularly from parents and teachers, in enhancing students' learning motivation (Banerjee & Halder, 2021). Positive psychology research emphasizes their supportive roles—emotional, academic, and resource-based—in boosting students' self-esteem, confidence, and motivation (Chirkov & Ryan, 2001). As foundational figures in education, parents and teachers play complementary roles in shaping students' academic outcomes and well-being, making their support key predictors of mathematics learning motivation (Banerjee & Halder, 2021; Sadoughi & Hejazi, 2023).

However, the mechanisms through which parental and teacher support influence students' internal psychological processes remain insufficiently understood. Research suggests that mathematics learning motivation is shaped not only by external support but also by internal factors such as growth mindset (Kaufmann, 2024) and self-efficacy (Schunk, 1995). Growth mindset, the belief that abilities are made possible through effort, fosters resilience in facing challenges (Dweck, 2006). Parental and teacher support can cultivate this mindset, enhancing students' coping abilities and learning motivation (Ha & Han, 2021; Ma et al., 2020). However, the interplay between external support, growth mindset, and motivation is underexplored. Self-efficacy, defined as confidence in task success, is another critical psychological factor influencing learning motivation (Schunk, 1995). Support from parents and teachers strengthens self-efficacy, which in turn boosts motivation and shapes learning behaviors (Zhao & Qin, 2021). PISA 2022 highlights the strong correlation between learning confidence and motivation, showing that higher self-efficacy leads to greater motivation (OECD, 2024). While some studies have examined the moderating role of external support (Bagci, 2018), the mediating role of self-efficacy in this context remains unclear. Furthermore, social theories of achievement motivation underscore the synergistic interaction between growth mindset and self-efficacy in influencing students' academic performance (Diseth et al., 2014; Dweck et al., 1995). However, how these factors jointly mediate the relationship between external support and learning motivation remains ambiguous.

Previous studies have largely examined external and internal influences on learning motivation separately, with limited focus on the interconnected pathways linking them. This study proposes that learning motivation, as a multifaceted construct, operates through dynamic interactions between external support (parental and teacher support) and internal psychological mechanisms (growth mindset and self-efficacy). To explore these relationships, the study addresses three questions:

- Q1. What is the current state of high school graduates' mathematics learning motivation?
- Q2. How do parental and teacher support influence mathematics learning motivation?
- Q3. What roles do growth mindset and self-efficacy play in mediating these effects?

A structural equation model (Figure 1) was developed to test nine hypothesized pathways connecting external support, internal mechanisms, and learning motivation.

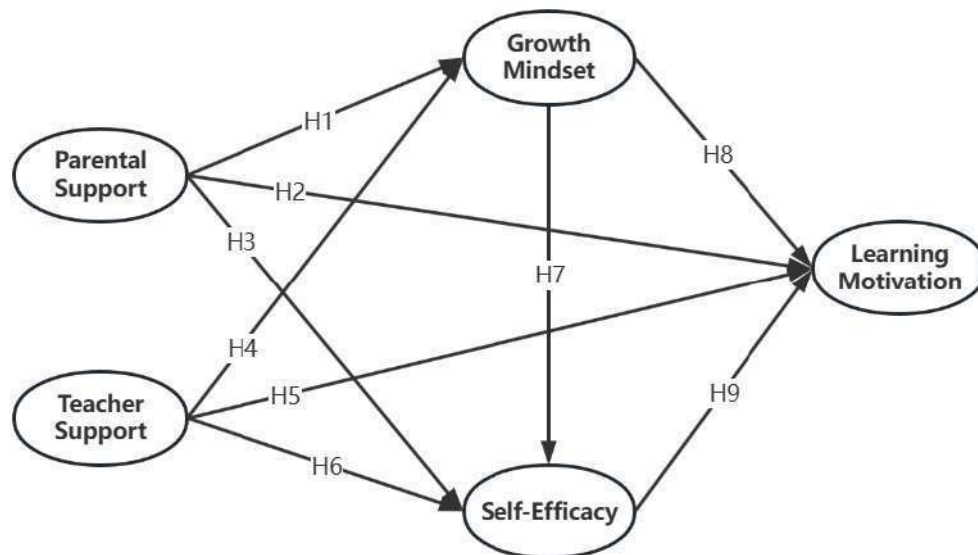


Figure 1. Research Model

2 METHODS

2.1 Participants

Considering that students who have completed the NCEE can objectively reflect the outcomes of foundational education, avoid interference from external evaluation systems to a certain extent, and provide valuable insights into the cultivation of mathematics learning motivation in higher education, this study selected Chinese post-examination students as the research participants. Data collection was conducted from June 20 to 30, 2024 (the NCEE in China is held annually on June 7 to 8). Using a convenient random sampling method, 336 valid samples were obtained, 37.20% were male and 62.80% were female.

2.1 Research Instruments

Before designing the questionnaire, focus group interviews were conducted to validate the rationality of the variables. Subsequently, relevant scales from domestic and international studies were collected and translated. Finally, experts in mathematics education reviewed the questionnaire's context to ensure its validity.

The finalized questionnaire comprises four modules. The first module introduces the purpose and scope of the study, defining key concepts. The second module collects demographic information. The third module forms the core of the study, assessing high school graduates' mathematics learning motivation and the underlying psychological mechanisms using a 5-point Likert scale. This module measures five latent variables, each represented by four items. Specifically, mathematics learning motivation focuses on students' drive to engage in mathematics learning, with statements such as "I enjoy learning mathematics" (Bagci, 2018). Growth mindset examines students' positive attitudes toward mathematics, with items like "Challenges in mathematics are opportunities for my growth and learning" (El-Adl & Alkharusi, 2020). Self-efficacy evaluates students' confidence in learning mathematics, including items such as "I

believe I can overcome difficulties in learning mathematics” (Schunk, 1995). Parental support reflects the role of parents in fostering mathematics learning, with statements such as “My parents/guardians care about my mathematics homework and academic progress” (Chirkov & Ryan, 2001). Teacher support assesses the assistance provided by teachers, with items such as “When I encounter difficulties in mathematics, my teacher helps me” (Chirkov & Ryan, 2001).

For data analysis, PLS-SEM was employed, as it is well-suited for exploring complex pathways, including multiple and chained mediators (Hair et al., 2011). Data analysis was conducted using Smart-PLS 4.0 software.

3 RESEARCH RESULT

3.1 Descriptive Statistics

As shown in Table 1, the standard deviations of individual items range from 0.958 (GM1) to 1.317 (SE1), while the mean scores range from 2.833 (SE1) to 3.929 (GM1) on a scale of 1 to 5. These results of Table 1 indicate that most high school graduates exhibit relatively high levels of mathematics learning motivation. They perceive greater support from teachers compared to parents, maintain an optimistic growth mindset toward mathematics learning, but demonstrate relatively moderate confidence in their mathematical abilities.

3.2 Measurement Model

The analysis of the measurement model primarily includes validity and reliability testing, shown in Tables 1 and 2.

Table 1. Reliability and Validity

Dimension	Item	Mean	SD	Factor Loading	CA	CR	AVE
Learning Motivation	LM1	3.551	1.135	0.909	0.900	0.909	0.769
	LM2	3.593	1.060	0.850			
	LM3	3.542	1.128	0.913			
	LM4	3.429	1.259	0.834			
Growth Mindset	GM1	3.929	0.958	0.769	0.849	0.863	0.688
	GM2	3.779	0.964	0.871			
	GM3	3.601	1.061	0.863			
	GM4	3.688	1.064	0.810			
Self-Efficacy	SE1	2.833	1.317	0.801	0.906	0.916	0.782
	SE2	3.316	1.087	0.924			
	SE3	3.423	1.069	0.928			
	SE4	3.628	1.067	0.878			
Parental Support	PS1	3.131	1.160	0.819	0.864	0.866	0.712
	PS2	3.376	1.152	0.897			
	PS3	3.565	1.044	0.850			
	PS4	3.539	1.079	0.807			
Teacher Support	TS1	3.926	1.025	0.866	0.927	0.931	0.820
	TS2	3.874	0.986	0.917			
	TS3	3.902	0.985	0.915			

TS4	3.777	1.015	0.922
-----	-------	-------	-------

Table 2. Discriminant Validity (Fornell-Larcker Criterion)

	LM	GM	SE	PS	TS
LM	0.877				
GM	0.617	0.829			
SE	0.851	0.628	0.884		
PS	0.456	0.416	0.523	0.844	
TS	0.537	0.568	0.561	0.537	0.905

In the measurement model of this study, Cronbach's Alpha (CA), Composite Reliability (CR), and factor loadings all exceeded 0.7, indicating good reliability. The Average Variance Extracted (AVE) values were all greater than 0.5, demonstrating satisfactory convergent validity. Additionally, the square roots of the AVE values for all latent variables were higher than their correlations with other latent variables, confirming adequate discriminant validity (Hair et al., 2011).

3.3 Structural Model and Hypothesis Testing

The validation of the structural model focuses on evaluating its predictive capability and the causal relationships between constructs. In this study, the R^2 value for mathematics learning motivation is 0.737, for self-efficacy is 0.498, and for growth mindset is 0.339. Higher R^2 values indicate stronger explanatory power, with thresholds for strong, moderate, and weak explanatory power being 0.67, 0.33, and 0.19, respectively (Hair et al., 2011). Therefore, the model demonstrates good explanatory power. Additionally, the Goodness of Fit (Gof) index is 0.63, exceeding the high standard threshold of 0.36, indicating an excellent model fit (Hair et al., 2011). Finally, the study tested the nine hypotheses proposed in the model. The structural model is illustrated in Figure 2.

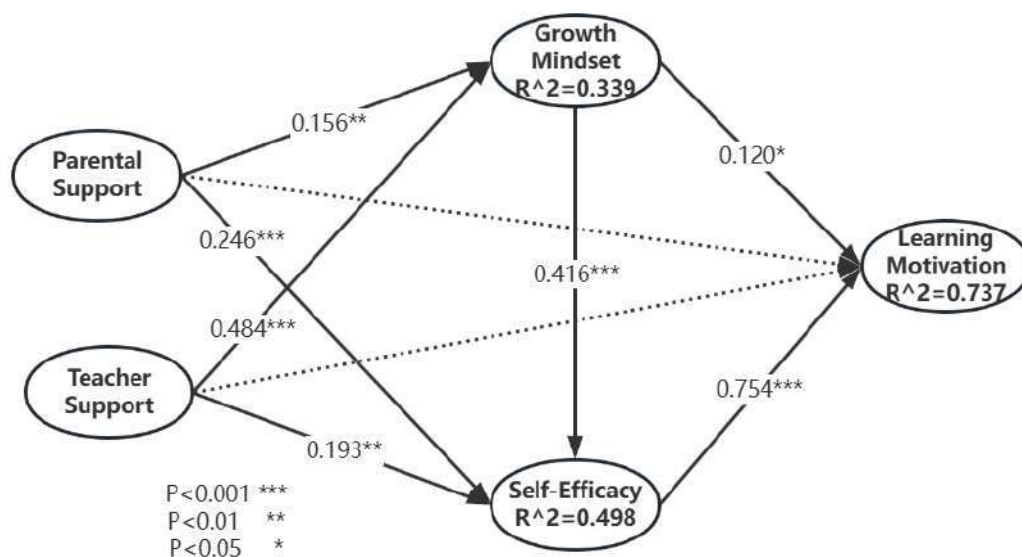


Figure 2. Structural model of this research

Furthermore, focusing on the effects of parental and teacher support on mathematics

learning motivation, five of mediating pathways were supported and identified as full mediation effects (see Table 3). This indicates that parental and teacher support have no direct influence on learning motivation but instead affect it through self-efficacy, growth mindset, or their chained mediation pathways.

Table 3. Mediating Effects

Path	β	M	SD	T	P	Mediation
TS -> GM -> LM	0.058	0.058	0.023	2.483	0.013	Full Mediation
TS -> SE -> LM	0.146	0.144	0.046	3.197	0.001	Full Mediation
TS -> GM -> SE -> LM	0.152	0.152	0.028	5.403	0.000	Full Mediation
PS -> GM -> LM	0.019	0.019	0.010	1.824	0.068	Not Significant
PS -> SE -> LM	0.186	0.187	0.042	4.378	0.000	Full Mediation
PS -> GM -> SE -> LM	0.049	0.049	0.019	2.600	0.009	Full Mediation

4 DISCUSSION

4.1 High School Graduates Exhibit Strong Mathematics Learning Motivation

Based on the average scores of the mathematics learning motivation dimension in the scale, it was found that high school graduates demonstrate strong mathematics learning motivation. This shows that positive attitude toward mathematics learning and a stable intrinsic drive, laying a solid foundation for their future academic development.

4.2 Parental and Teacher Support Influence Students' Mathematics Learning Motivation Indirectly

Previous studies have revealed that support from significant others exerts a direct influence on students' learning motivation (Banerjee & Halder, 2021; Chirkov & Ryan, 2001). However, hypothesis testing in this study indicates that in the context of mathematics, the direct influence of parental and teacher support on high school graduates' learning motivation is not significant. Instead, their influence is mediated through the interaction with students' internal psychological factors, producing an indirect effect. This conclusion is also supported by other research findings (Zhao & Qin, 2021). It is recommended that efforts to enhance students' learning motivation should not solely focus on external support but also emphasize the development of students' internal psychological mechanisms. A holistic approach integrating external support and internal psychological factors is crucial.

4.3 Growth Mindset and Self-Efficacy as Mediators

Firstly, growth mindset fully mediates the relationship between teacher support and mathematics learning motivation. Greater perceived teacher support fosters a growth mindset, which enhances learning motivation (Sadoughi & Hejazi, 2023). Teachers should emphasize effort when students face challenges to develop their growth mindset. Secondly, self-efficacy mediates the effects of parental and teacher support on learning motivation. Higher support strengthens self-efficacy, which boosts motivation (Schunk, 1995). Parents and teachers should collaborate to create a supportive environment that enhances self-efficacy and motivation (Chirkov & Ryan, 2001).

Finally, growth mindset and self-efficacy jointly exhibit a chained mediation effect. Growth mindset fosters positive beliefs about challenges, while self-efficacy enhances motivation, together with driving learning motivation (Dweck et al., 1995). Educators should foster an understanding of ability growth and build confidence through incremental successes.

5 IMPLICATIONS

This study analyzed the synergistic pathways of internal and external drivers of learning motivation, yet some limitations remain. First, the research employed a cross-sectional design; future studies could use longitudinal approaches to explore the evolutionary mechanisms of learning motivation. Second, the sample coverage was relatively limited, which may affect the generalizability of the findings. Third, this study primarily focused on the roles of parental and teacher support; future research could investigate the influence of other significant individuals.

References

- Aeschlimann, B., Herzog, W., & Makarova, E. (2016). How to foster students' motivation in mathematics and science classes and promote students' STEM career choice. A study in Swiss high schools. *International Journal of Educational Research*, 79, 31–41.
- Bagci, S. C. (2018). Does Everyone Benefit Equally From Self-Efficacy Beliefs? The Moderating Role of Perceived Social Support on Motivation. *The Journal of Early Adolescence*, 38(2), 204–219.
- Banerjee, R., & Halder, S. (2021). Effect of Teacher and Parent Autonomy Support on Academic Motivation: A Central Focus of Self-Determination Theory. *World Futures*, 77(6), 452–480.
- Chirkov, V. I., & Ryan, R. M. (2001). Parent and Teacher Autonomy-Support in Russian and U.S. Adolescents: Common Effects on Well-Being and Academic Motivation. *Journal of Cross-Cultural Psychology*, 32(5), 618–635.
- Deci, E. L., & Ryan, R. M. (2000). The 'What' and 'Why' of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry*, 11(4), 227–268.
- Diseth, Å., Meland, E., & Bredablik, H. J. (2014). Self-beliefs among students: Grade level and gender differences in self-esteem, self-efficacy and implicit theories of intelligence. *Learning and Individual Differences*, 35, 1–8.
- Dweck, C. S. (2006). *Mindset: The New Psychology of Success*. Random House Publishing Group.
- Dweck, C. S., Chiu, C., & Hong, Y. (1995). Implicit Theories and Their Role in Judgments and Reactions: A Word From Two Perspectives. *Psychological Inquiry*, 6(4), 267–285.
- El-Adl, A., & Alkharusi, H. (2020). Relationships between Self-Regulated Learning Strategies, Learning Motivation and Mathematics Achievement. *Cypriot Journal of Educational Sciences*, 15(1), 104–111.
- Ha, J.-H., & Han, C. (2021). Converged Study of Perceived Parental Autonomy Support, Growth Mindset, Grit, and Help-Seeking Behaviors Of High-School Students. *Journal of the Korea Convergence Society*, 12(6), 161–171.
- Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a Silver Bullet. *Journal of Marketing Theory and Practice*, 19(2), 139–152.

- Kaufmann A. (2024). Investigating the Relationship between a Growth Mindset and Intrinsic Motivation in a Cognitive Task. University of Twente.
- Ma, C., Ma, Y., & Lan, X. (2020). A Structural Equation Model of Perceived Autonomy Support and Growth Mindset in Undergraduate Students: The Mediating Role of Sense of Coherence. *Frontiers in Psychology, 11*.
- OECD. (2024). PISA 2022 Results (Volume V): Learning Strategies and Attitudes for Life, PISA. OECD Publishing, Paris. <https://doi.org/10.1787/c2e44201-en>
- Sadoughi, M., & Hejazi, S. Y. (2023). The effect of teacher support on academic engagement: The serial mediation of learning experience and motivated learning behavior. *Current Psychology, 42*(22), 18858–18869.
- Schunk, D. H. (1995). Self-efficacy, motivation, and performance. *Journal of Applied Sport Psychology, 7*(2), 112–137.
- Suharnadi, P., S, N., & Nirwana, H. (2024). The Role and Function of Learning Motivation in Improving Student Academic Achievement. *Manajia: Journal of Education and Management, 2*(1), Article 1.
- Zhao, J., & Qin, Y. (2021). Perceived Teacher Autonomy Support and Students' Deep Learning: The Mediating Role of Self-Efficacy and the Moderating Role of Perceived Peer Support. *Frontiers in Psychology, 12*.

SECONDARY STUDENTS' UNDERSTANDING OF THE FUNCTIONS OF PROOF FOR THE EQUIVALENT MEANINGS OF GENERAL VALIDITY

Sarah Lundt, Milena Damrau, and Stefan Ufer

LMU Munich, Germany

The validity of a universal statement is equivalent to the inexistence of counterexamples. One function of proof is to establish both of them. We investigate to which extent students show consistent convictions regarding the validity of a statement, the inexistence of counterexamples, and the validity of a given proof. Our results indicate that a substantial proportion of students do not only have difficulties in understanding the equivalence between a statement's validity and the inexistence of counterexamples but also in understanding the respective functions of proof. Students seem to not systematically draw on general principles connected to the functions of proof but rather on the mathematical content of the statement.

INTRODUCTION

One central characteristic of mathematics is that there are statements that apply to all, possibly infinitely many, elements of a certain set without any exception. Those statements are called *universal statements* and are often of the form 'For all objects o within a given set S , property $P(o)$ holds.' (Damrau, 2023). Since valid universal statements apply to *all* elements o of the given set S , there are no counterexamples to these statements, i.e., no objects o of the set S for which $P(o)$ does not hold. This also means that a universal statement is invalid if just one counterexample exists. Clearly, the validity of a universal statement is logically equivalent to the inexistence of counterexamples. We refer to this relation as *equivalent meanings of general validity*.

Students' understanding of the equivalent meanings of general validity

The crucial role of universal statements and their proofs in mathematics is mirrored in educational standards worldwide (e.g., KMK, 2012). Although secondary school students are expected to learn to deal with (universal) statements and their proofs, numerous studies showed that these topics pose considerable difficulties for many learners (e.g., Healy & Hoyles, 2000; Sommerhoff et al., in press). Many of these studies focused on proof construction, among others identifying a lack of knowledge of mathematical concepts (Sommerhoff et al., in press) as possible reasons for difficulties. Furthermore, students seem to have difficulties understanding that the validity of universal statements and the inexistence of counterexamples are equivalent. Chazan (1993) presents cases of students, who were absolutely convinced of the validity of a universal statement were not absolutely convinced that there is no counterexample. Barkai et al. (2002) reported that many learners did not accept one

single counterexample as sufficient to refute the validity of a universal statement. Damrau (2023) found that in about a third of all observations, first-year university students had differently strong convictions of the validity of a universal statement and the inexistence of counterexamples. A quarter of the investigated students even stated explicitly, that a universal statement must be seen true in many cases or generally true, but potentially with a few exceptions. Variability in these difficulties was related to student characteristics and characteristics of the universal statements.

Students' understanding of the functions of proof

One educational approach to address the difficulties regarding general validity would be to draw on the function of proof to establish the validity of universal statements and thereby excluding the existence of counterexamples. Already Fischbein and Kedem (1982) argued, that many students do not understand that a proof verifies the validity of a universal statement (*verification function of proof*, de Villiers, 1990). Furthermore, Shongwe (2021) described interindividual differences in the understanding of different proof functions, including the verification function, in grade 11 learners.

To prove that a universal statement is valid, usually a particular but arbitrary (*generic*) object of the set S is chosen and represented by a *variable* symbol (e.g., x) (Durand-Guerrier, 2017). A valid proof then deduces that all objects of S referenced by x have property $P(x)$, using accepted rules of inference and drawing on accepted statements from a framing theory, which hold for all objects o of S (Stylianides, 2007). Using variables to refer to the arbitrary object x indicates that the proof applies to all objects of the set S . Since no specific properties of a specific object of S are used, but only properties that hold for all objects of S , the proof establishes the validity of the universal statement for all objects of S .

By drawing on variables, a proof simultaneously excludes the existence of counterexamples: Potential counterexamples, i.e., specific objects c of the set S for which it is uncertain whether $P(c)$ holds, can be replaced for the variable x in the proof. This results in a very specific proof showing that property $P(c)$ also holds for this particular object c . Therefore, c cannot be a counterexample. Logically, this *excluding counterexamples function of proof* is equivalent to the verification function. Due to learners' problems grasping this equivalence, we distinguish these two functions. We summarize these concepts and relations in the conceptual framework shown in Figure 1. Note that the validity of the proof implies the validity of the universal statement and the inexistence of counterexamples, but not vice versa.

Students' problems to understand the equivalent meanings of general validity (Damrau, 2023), raise the question to which extent students understand that (and how) proofs support the validity of universal statements and the inexistence of counterexamples. Moreover, understanding these two functions of proof might be hindered by limited understanding of the equivalent meanings of general validity.

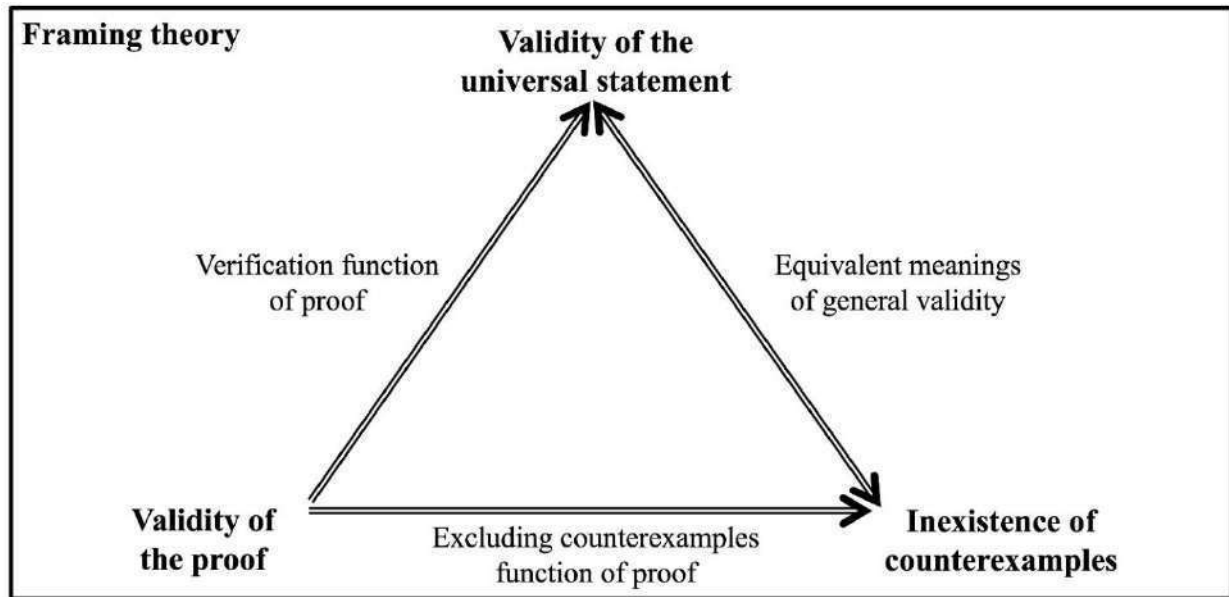


Figure 1: Conceptual framework

GOALS OF THE PRESENT STUDY

The present study thus addresses the following questions: (Q1) To which extent do secondary school students show consistent convictions regarding a) the validity of a universal statement, b) the inexistence of counterexamples, and c) the validity of a given proof? (Q2) Do the results for Q1 differ for two universal statements of different familiarity? (Q3) To which extent do the students refer to a given proof when justifying the validity of a universal statement or the inexistence of counterexamples?

METHODS

We conducted a 60-minute cross-sectional computer-based in-person questionnaire study with a sample of $N = 78$ grade eight students from the high attaining school track in Germany. We first provided students with a recap of the framing theory. Students then were randomly assigned to work initially on one of two universal statements: (A) “The sum of *any* two even numbers is *always* even.”, (B) “For *every* natural number n , the value of the term $n^3 - n$ is *always* divisible by 6.”. We consider statement A to be more familiar and easier accessible in terms of content and constructing or investigating examples. Roughly one third of the students also completed work on the other statement, after finishing the first one. Each time, students first read the universal statement and a corresponding variable-based proof, without receiving any information about whether the statement or the proof are valid. Afterwards, students reported their conviction regarding the validity of the universal statement, the (in)existence of counterexamples and the validity of the proof (Fig. 1) on a four-point scale (absolute and relative conviction for or against it, Weber & Mejía Ramos, 2015) and were asked to provide justifications for their answers in writing.

Participants’ justifications were coded based on Harel and Sowder (1998) and Damrau (2023) ($\kappa = 0.90$ for about 20% of the data). Answers that do not contain a discernible

justification were coded as *deficient*. Answers that refer to the fact that the statement has already been identified as valid by an external source (e.g., by a teacher) were coded as *external*. Repetitions or paraphrases of the statement were coded as *pseudo-arguments*. Example-based arguments were coded as *empirical*, with counterexamples coded separately as (*supposed*) *counterexample*. If students (tried to) construct their own arguments, these were coded as (*own*) *transformational argument* or (*own*) *deductive argument*. Rough proof attempts also fall into these categories. Answers that referred to the given proof in any way were coded as *related to the given proof*.

RESULTS

Descriptive results on students' convictions

Overall, a substantial number of students was absolutely convinced that statement A is valid (65%), that no counterexample exists (77%), and that the proof is valid (49%). For statement B, a large amount of students expressed relative conviction regarding the statement's validity (49%) and the inexistence of counterexamples (40%). About one third of the students were relatively convinced that the proof for statement B is valid, and roughly equally many were relatively convinced that it is invalid.

The statement's validity and the inexistence of counterexamples (Q1, Q2)

Regarding statement A, most of the students (63%) were equally convinced of the validity of the statement and the inexistence of counterexamples. The remaining students were mostly more convinced of the inexistence of counterexamples than of the validity of the statement (23%). In contrast, fewer students (50%) were equally convinced of both for statement B, and most of the remaining students (29%) reported stronger conviction regarding the validity of the statement than the inexistence of counterexamples.

The proof's and the statement's validity (Q1, Q2)

Both statements yielded similar result patterns regarding students' convictions of the validity of the universal statement and its proof. About half of the students were equally convinced of the validity of the proof and the universal statement, while about one third of them were more convinced of the validity of the statement. Both are consistent with the verification function of proof, but the latter indicates that students might draw on alternative reasons, apart from the given proof, for believing in the statement's validity. The remaining students (~15%) were more convinced of the validity of the proof than of the statement, indicating that they might have not fully grasped the verification function of proof. This holds in particular for 6 students (~10%) for statement A and for 4 students (~8%) for statement B who were absolutely convinced of the validity of the proof, but at most relatively convinced of the validity of the statement.

The proof's validity and the inexistence of counterexamples (Q1, Q2)

For statement A, about half of the students were equally convinced of the validity of the proof and the inexistence of counterexamples. Just under 40% of the students

showed stronger convictions towards the inexistence of counterexamples. Again, both are consistent with the underlying function of proof, but the latter again indicates that students might draw on alternative reasons, apart from the given proof, for believing that no counterexamples exist. The remaining 10% were more convinced of the validity of the proof, again indicating that they might not have fully grasped the corresponding function of proof. For statement B, 42% of the students were equally convinced, while 25% were more convinced of the inexistence of counterexamples, again indicating alternative reason to believe in the inexistence of counterexamples. The remaining third was more convinced of the validity of the proof than of the inexistence of counterexamples, indicating an incomplete understanding of the function of proof to exclude the existence of counterexamples.

Fisher's Exact Test yielded that students' convictions regarding the statement's validity, the inexistence of counterexamples, and the proof's validity are significantly related to each other for statement B (all p -values $< .05$), but not for statement A.

Justifications for statement validity and inexistence of counterexamples (Q3)

For statement A 70% and for statement B 72% of the students justified their rating of the statements' validity. Regarding the inexistence of counterexamples, 81% justified their answer for statement A and 66% for statement B. The categories *External* and *(Own) transformational argument* were not observed in the data. The frequencies for the remaining categories (relative to all possible justifications) are presented in Fig. 2.

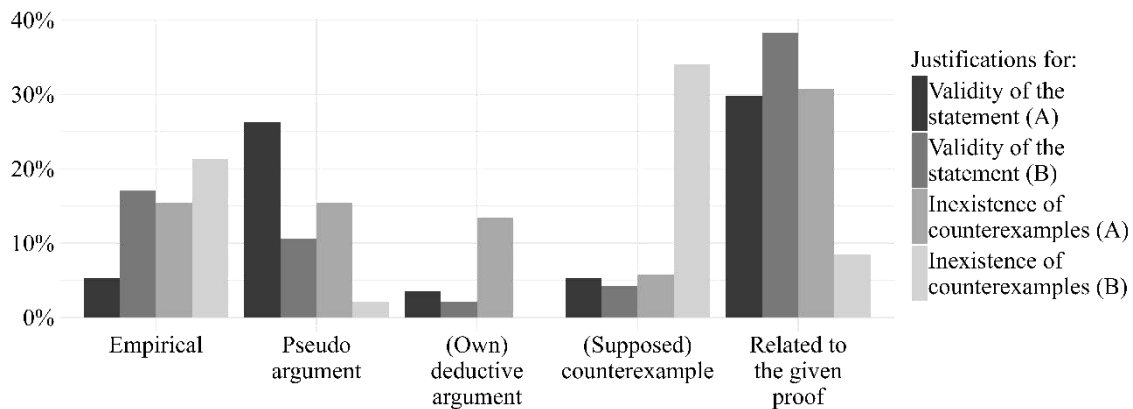


Figure 2: Justifications for statement validity and inexistence of counterexamples

To justify the statement's validity, only few students presented own deductive arguments. Empirical arguments, such as giving supporting examples, were rather rare for both statements (A: 5%, B: 17%). Arguments that went beyond pseudo-arguments such as repeating the statement itself were mostly related to the given proof (A: 30%, B: 38%). These proof-related arguments were also made by students who were convinced that the proof was invalid. Considering these justifications more closely, these students justified the (supposed) invalidity of the statement with the (supposed) invalidity of the proof (e.g., because they thought there is a mistake in the proof). Moreover, by far not all of the students who were absolutely convinced of the validity

of the statement and the validity of the proof justified the validity of the statement with the validity of the proof (A: 33%, B: 60%).

In contrast, some students (A: 15%, B: 21%) used empirical arguments to justify their conviction regarding the inexistence of counterexamples. For statement A, most students, again, referred to the given proof. For statement B, only few justifications were related to the given proof (9%). Again, by far not all of the students who were absolutely convinced of the inexistence of counterexamples and the validity of the proof justified the inexistence of counterexamples with the validity of the proof (A: 50%, B: 20%). For statement B, (supposed) counterexamples were provided quite frequently (34%), which were, of course, incorrect (e.g., assuming that 0 is not divisible by 6). There were also students who provided (supposed) counterexamples even though they considered the proof to be valid. This indicates that not understanding the functions of proof may inhibit students' chances to reconsider and develop insecure knowledge of the statements' mathematical content (e.g., divisors of 0).

DISCUSSION

A substantial proportion of students seem to be insecure about the equivalence of a universal statement's validity and the inexistence of counterexamples (Q1), and this proportion is larger for secondary school students in our study (37% to 50%) than for first-year university students (roughly 30%, on average) reported in Damrau (2023). Our findings further connect Damrau's work on the general validity of universal statements with works on the understanding of the functions of proof in mathematics (e.g., Shongwe, 2021). Secondary school students seem to have problems understanding the functions of proof to establish the statement's validity and the inexistence of counterexamples, as a number of students was more convinced of the proof's validity than of the statement's validity or the inexistence of counterexamples.

This result pattern regarding students' conviction is particularly strong for statement B (Q2) again resonating with results of Damrau (2023): Students showed more consistent convictions for the more familiar statement, which could possibly be attributed to the generally strong positive convictions towards familiar statement A. The results of Fisher's Exact Test support this assumption, as we found no evidence that the different convictions towards statement A are significantly related to each other. This could also explain why the result pattern for Q1 is particularly strong for statement B, since the results for statement A might be masked by students' strong positive convictions.

The results for Q1 and Q2 from students' (closed) ratings are also mirrored in and complemented by students' written (open) justifications. Firstly, there were students, who justified the (supposed) invalidity of the statement with the (supposed) invalidity of the proof. These students may have applied an (invalid) converse of the verification function of proof: An invalid proof invalidates the statement. Secondly, some of the justifications indicate that some students seem to draw on alternative sources of evidence, other than the proof, to conclude that the statement is valid. These alternative

sources of evidence may relate to specific examples (which were, however, rarely mentioned, in particular for statement A) or mere prior experience with the statement. The justifications for the function of proof to exclude counterexamples yielded a similar pattern of results. Notably, these similar results occur, even though a sizeable number of students does not seem to understand the equivalence of the statement's validity and the inexistence of counterexamples. Again, not all of the students who were absolutely convinced of the validity of the proof use this as justification to exclude counterexamples. While this may be explained by the strong positive convictions towards the familiar statement A, this explanation would not hold for statement B. For statement B, a substantial proportion of students were convinced that counterexamples exist and even named concrete (supposed and incorrect) counterexamples, often in spite of being convinced of the proof's validity.

Adding to and combining results of Damrau (2023) and Shongwe (2021), our findings indicate that secondary school students do not only show problems in understanding the general validity of a universal statement itself, but also regarding the functions of proof for this general validity. Most students, including those convinced of the proof's validity, do not systematically draw on general principles connected to the functions of proof. Moreover, some answers obviously mirror misunderstandings regarding such general principles. We further find that students draw on alternative evidence rather than the given proof to evaluate the validity of a statement or the inexistence of counterexamples, most likely based on their understanding of the statements' contents. Generally, the justifications indicate that—potentially independent of the assumed validity of the proof—students seem to draw more on their (sometimes partial) understanding of the statements' mathematical content to evaluate the statements' validity as well as the inexistence of counterexamples, which they do not identify as equivalent (cf. Q1). This resonates with the role of knowledge of mathematical contents for proof-related performance (Sommerhoff et al., in press). It also explains the supporting role of statement A's familiarity and the problems with insecure mathematical knowledge in statement B. This reliance on the statements' contents alone is indicative of an insufficient understanding of the functions of proof. Even more, this insufficient understanding potentially inhibits students' chances to identify own insecure understanding of the mathematical contents.

While the reliance on the statements' contents may not be a problem on itself, a possible explanation might be either that students do not understand the functions of proof, that they are not able to identify the proofs as valid (proof validation), or that they do not comprehend the given proofs (proof comprehension, Mejía Ramos et al., 2012) to an extent that allows them to see the proofs' potential to answer questions on the statements' validity and/or the inexistence of counterexamples. More precisely, proof comprehension may specifically relate to the role of variables for establishing the general validity of universal statements. Since little evidence was found for students explicitly drawing on the functions of proof, experimental scaffolding studies

explaining the way mathematics establishes general validity and eliciting or supporting more principle-based ways of reasoning may be promising.

References

- Barkai, R., Tsamir, P., Tirosh, D., & Dreyfus, T. (2002). Proving or refuting arithmetic claims: The case of elementary school teachers. In A. Cockburn & E. Nardi (Eds.), *Proceedings of the 26th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 57-64). PME.
- Chazan, D. (1993). High school geometry students' justification for their views of empirical evidence and mathematical proof. *Educational Studies in Mathematics*, 24(4), 359-387.
- Damrau, M. (2023). Understanding the generality of mathematical statements: An experimental study at the transition from school to university. Springer
- de Villiers, M. (1990). The role and function of proof in mathematics. *Pythagoras*, 24, 17-24.
- Durand-Guerrier, V. (2017). Use of letters in mathematics at university level teachers' practices and students' difficulties. In R. Göller, R. Biehler, R. Hochmuth, & H.-G. Rück (Eds.), *Didactics of Mathematics in Higher Education as a Scientific Discipline* (pp. 106-110). Universitätsbibliothek Kassel.
- Fischbein, E., & Kedem, I. (1982). Proof and certitude in the development of mathematical thinking. In A. Vermandel (Ed.), *Proceedings of the 6th Conference of the International Group for the Psychology of Mathematics Education* (pp. 128-131). PME.
- Harel, G., & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. In A. H. Schoenfeld, J. Kaput, & E. Dubinsky (Eds.), *Research in Collegiate Mathematics Education III* (Vol. 7, pp. 234-283). American Mathematical Society.
- Healy, L., & Hoyles, C. (2000). A study of proof conceptions in algebra. *Journal for Research in Mathematics Education*, 31(4), 396-428.
- KMK. (2012). Bildungsstandards im Fach Mathematik für die Allgemeine Hochschulreife: Beschluss der Kultusministerkonferenz vom 18.10.2012.
- Mejía Ramos, J. P., Fuller, E., Weber, K., Rhoads, K., & Samkoff, A. (2012). An assessment model for proof comprehension in undergraduate mathematics. *Educational Studies in Mathematics*, 79(1), 3-18.
- Shongwe, B. (2021). Learners' beliefs about the functions of proof: building an argument for validity. *Educational Studies in Mathematics*, 107(3), 503-523.
- Sommerhoff, D., Kollar, I., & Ufer, S. (in press). Undergraduate students' mathematical proof skills: Examining the impact of six cognitive re-sources. In K. Weber & M. Savic (Eds.): *New Directions for Mathematics Education Research on Proving: Honoring the Legacy of John and Annie Selden*. Springer.
- Stylianides, A. J. (2007). Proof and proving in school mathematics. *Journal for Research in Mathematics Education*, 38(3), 289-321.
- Weber, K., & Mejía Ramos, J. P. (2015). On relative and absolute conviction in mathematics. *For the Learning of Mathematics*, 35(2), 15-21.

EMBODIED ARGUMENTATION: CONVINCING YOURSELF, YOUR FRIEND, OR THE SKEPTIC?

Nadav Marco^{1,2} and Alik Palatnik¹

¹The Hebrew University of Jerusalem, ²David Yellin Academic College of Education

An essential function of proofs is to establish conviction in the truth of mathematical statements. However, formal proofs do not always yield personal or interpersonal conviction. This design-based research explores how students can construct convincing proofs through spatial geometry activities with tangible, body-sized models. We identify a phenomenon called embodied argumentation (EA)—the use of bodily movements and interactions with physical objects to construct and present purported proofs. Focusing on an undergraduate student working on the Midsegments in a Cube Task, we examine if EA served to convince (1) oneself, (2) others, and (3) a skeptic. Our findings show that EA can fulfill all these roles, showing its potential to bridge formal proving and conviction.

Recent research in mathematics education highlights the potential of embodied learning, particularly in geometry (Palatnik et al., 2023). Grounded in 4E theories of cognition (embodied, enactive, embedded, and extended), embodied approaches emphasize how physical interaction and sensory experience support mathematical understanding (Abrahamson & Sánchez-García, 2016; Marco & Shvarts, 2025). Geometry, relying on spatial reasoning, is especially suited to such approaches. Mathematical argumentation, an essential purpose of geometry education, builds conviction among learners through justified statements (Herbst et al., 2017). Prior studies demonstrate how gesturing and using spatial models help learners connect abstract ideas with perceptual experiences (Nathan & Martinez, 2015; Ng et al., 2020; Walkington et al., 2019).

In previous studies, we explored Proofs Without Words as a visual resource to support students' formal proof construction (Marco et al., 2022). This paper builds on the concept of *embodied argumentation* (Palatnik et al., 2025), which reframes proof and reasoning by recognizing multimodal actions as integral to mathematical argumentation. Embodied argumentation is defined as “a form of mathematical argumentation in which learners use physical actions, such as interactions with objects and gestures, to construct, support, and communicate proofs” (p.1). Our previous work showed how body-scale geometric models scaffold reasoning (Palatnik et al., 2025). However, the role of embodied argumentation in bridging intuition and formal proof remains underexplored. This study investigates how it fosters personal or interpersonal mathematical conviction.

THEORETICAL UNDERPINNINGS AND RESEARCH QUESTION

Stylianides et al. (2017) identified three perspectives on proving: as problem-solving, as convincing, and as a socially embedded activity. This study aligns primarily with the second perspective, focusing on understanding students' or teachers' standards of mathematical conviction and their relation to disciplinary norms. Still, recognizing that conviction is socially constructed, we adopt Mason et al.'s (2010) framework, which emphasizes convincing *oneself*, a *friend*, and a *skeptic*. According to Mason et al., personal conviction involves intuitive reasoning to build confidence in a solution. Convincing a friend requires clear articulation and effective representations, often uncovering gaps in reasoning. Convincing a skeptic demands rigorous, generalizable arguments that withstand critique. Mason et al. stress the role of an "internal skeptic" in refining arguments, highlighting proof as an iterative process of personal and social validation. Guided by this framework, our study investigates the following research question: *How does embodied argumentation function in spatial-geometry proving, specifically, are embodied arguments used to convince oneself, a friend, or a skeptic?*

METHOD

This research is part of a long-term design-based research study (Bakker, 2018). Building on our identification of embodied argumentation during group work on the Icosahedron Task (Palatnik et al., 2025), our next iteration goal was to create a task that (a) would still encourage participants to engage with a geometric object in meso-space (Herbst et al., 2017), facilitating interactions such as rotations and adding auxiliary constructions and (b) will focus on an object more familiar in school curricula. We hypothesize that the Midsegments in a Cube Task (Figure 1) will foster embodied argumentation by engaging students in spatial reasoning. This design experiment aimed to evaluate this hypothesis and investigate our research question regarding the conviction of oneself, a friend, or a skeptic (Mason et al., 2010).

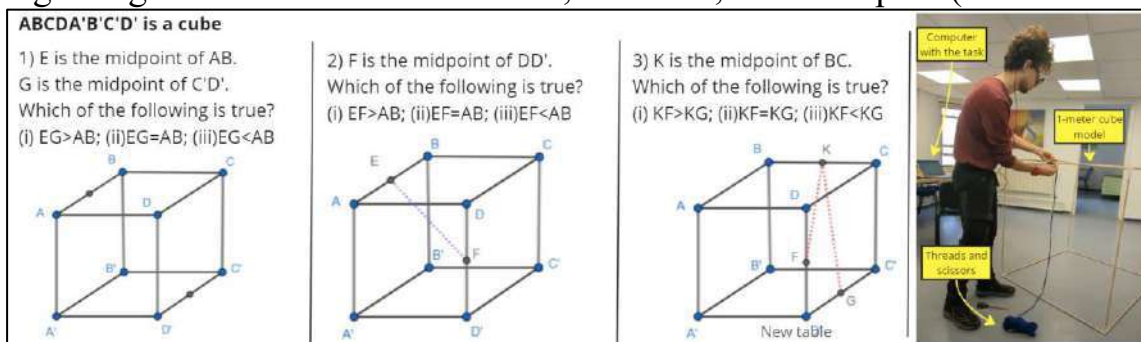


Figure 1: The Midsegments in a Cube Task and research settings.

Research settings

The task was presented on a computer positioned on a table against one wall of the research lab. At the center of the room stood a 1-meter cube constructed from wooden sticks and plastic joints (Figure 1). Blue fabric thread and scissors were provided near

the cube, allowing the participant to use and cut the thread as needed during the task. The participant, Edi, was instructed to explicate the arguments that convinced him, a friend, and a skeptic for each item. Additionally, he was asked to record in the answer-form the arguments he believed would persuade the most demanding skeptic. Edi recently completed his undergraduate degree in mathematics and physics and earned a teaching certificate in secondary mathematics. We invited Edi to participate due to his strong problem-solving abilities and familiarity with thinking aloud, as observed during our teaching interactions in his certification program.

Data collection and analysis

The session, lasting 18 minutes, was video-recorded using two cameras: a mobile eye-tracking device (Neon by Pupil Labs) worn by Edi and a portable side camera. Six days later, a 55-minute simulated recall interview was conducted and recorded. Edi watched the eye-tracking footage and provided insights, pausing at points he deemed significant to explain his reasoning and actions.

Data analysis followed a micro-ethnographic methodology (Streeck & Mehus, 2005) enhanced by eye-tracking (see Marco & Shvarts, 2025). The analysis proceeded in two stages. First, video recordings were reviewed iteratively to identify episodes that met our definition of embodied argumentation. For instance, rotating the cube to make two segments appear symmetrical. Second, these episodes were analyzed to determine whether the embodied arguments were used to convince oneself, a friend, or a skeptic. Verbal expressions of conviction were triangulated with data from the simulated recall and Edi's written responses in the answers form he completed during the task.

FINDINGS

In Edi's work, we found several salient cases of embodied argumentation. In this paper, we analyze two of them. We will show that the two cases satisfy our definition of embodied argumentation and argue that while Edi considered the first embodied argument convincing only to himself, he considered the second embodied argument convincing both for a friend and a skeptic.

Embodied argument for convincing only oneself

We join Edi after he correctly solved the first item by constructing EG and the height from E to A'B' (see Figure 2). When Edi read the second item aloud, he stopped at the description of EF, and before reading the question, he went to the model to construct it. While tying EF, he said: "Another thing I can do, if I want to be anti-rigorous and anti-mathematical, is to unravel these threads I've cut and compare their lengths to each other." This approach to proving could be considered a form of embodied argumentation. However, Edi's dismissal of its acceptability, along with his decision not to perform it, suggests that he would not find it convincing even for himself.

After Edi finished constructing the segment EF, he looked at the screen and read the question. Then he went to hold EF in both hands and slid back his left hand from E

towards A while keeping the right hand on F (Figure 2, 7:39). He let go of EF, stepped backward, and said: “So...” he remained silent for about seven seconds in which he tilted his head and gazed at FD (Figure 2, 7:43 and 7:46). He continued:

(7:50) It [EF] again should be longer because I have two projections here. I have here one projection in this direction [traces from E to A, Figure 2, 7:53] that shortens the side for me, and one projection in this direction [traces from F to D, Figure 2, 7:57] that shortens the side for me... This is why it [EF] is longer than the cube's side [gazing along AF].

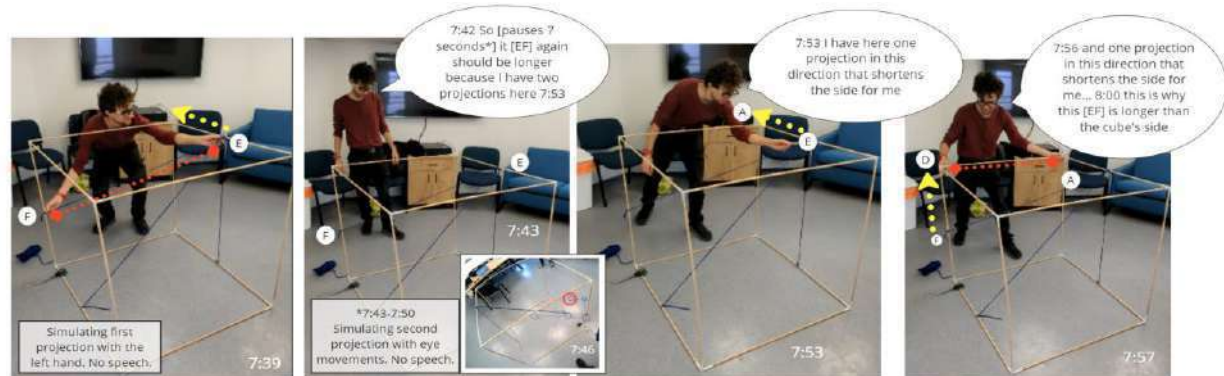


Figure 2: Embodied argumentation for proving that $EF > AD$

Analysis. Edi's actions and speech constitute a case of embodied argumentation because his claims were rooted in physical interaction and movement on the physical manipulative. When holding EF with his hands (Figure 2, left red mark), Edi could feel its length in tactile sensory-motor experience and his actions of projection (on the cube's faces) enabled him to feel this length gradually shortening until becoming AD (Figure 2, right red mark). His speech also hints at this proprioceptive feeling when he says twice that the projections shorten the side “for me.” We now turn to show that Edi considered this embodied argument only to convince himself.

Data. While Edi concluded that “[EF] is longer than the cube's side” (8:02), his eyes repeatedly fixated between A and F, suggesting a simulation of AF construction.

Interviewer 1: (8:06) So again, we return to the story of convincing a friend...

Edi: So, for a friend, I would construct [goes to pick up the thread and scissors]

Interviewer 2: Did what you just did [the projections] convince you?

Edi: It allows me to... to... yes, it convinces me because what I just did now with the projection motion... shortens distances in this context.

Next, Edi developed an alternative embodied argument by constructing an auxiliary line (AF) with a thread, which he later wrote down in the form to convince a skeptic:

It is possible to construct the side AF, which will be the hypotenuse of the right triangle AFD and, therefore, longer than a side of the cube. The triangle AEF is also a right triangle

because AB is perpendicular to the $A'D'D$ plan, and therefore, EF is the hypotenuse of a right triangle. Thus, $EF > AF > AD$.

In the stimulated recall, when asked about his decision to present a different argument to a friend Edi explained: "Projection is hard to see... figuring out which segment is projected onto which plane... constructing a triangle is a way to explain clearly to a friend that these things are true... I wanted to build the triangle that connects the two projections."

Analysis. Although Edi admitted that the first embodied argument with the projections convinced him, he did not use it to convince a friend or a skeptic. He perceived this argument as "hard to see"—without physically experiencing the movements, it is challenging to be convinced why the projection shortens the segment in this context. Moreover, Edi further emphasized the challenge of explaining the projection movement verbally to others ("figuring out which segment is projected onto which plane"). Nevertheless, Edi did not abandon the projection argument. Instead, he used it as inspiration to construct the auxiliary line "that connects the two projections." For him, this approach could convince a friend because it makes the essence of the projection argument more perceptually accessible ("a way to explain clearly").

Embodied argument for convincing a friend and a skeptic

Already when reading the third item on the screen, Edi cautiously conjectured that KG and KF would be equal "from symmetry (13:18)." He then connected a thread to K and performed a 90 degrees rotation of the cube on the $A'B'$ axis (Figure 3, 13:57-59) while saying (for the co-occurrence of speech and gestures see Figure 3):

The reason I maintain that they [KG and KF] are equal is that if I rotate the cube [goes to stand near the upright BC , facing $A'D'$] and bisect it along this axis [$CBA'D'$] so, I have symmetry. If I go from the midpoint of this side [K] and go here [G] or here [F], these two things are supposed to be at equal distances.

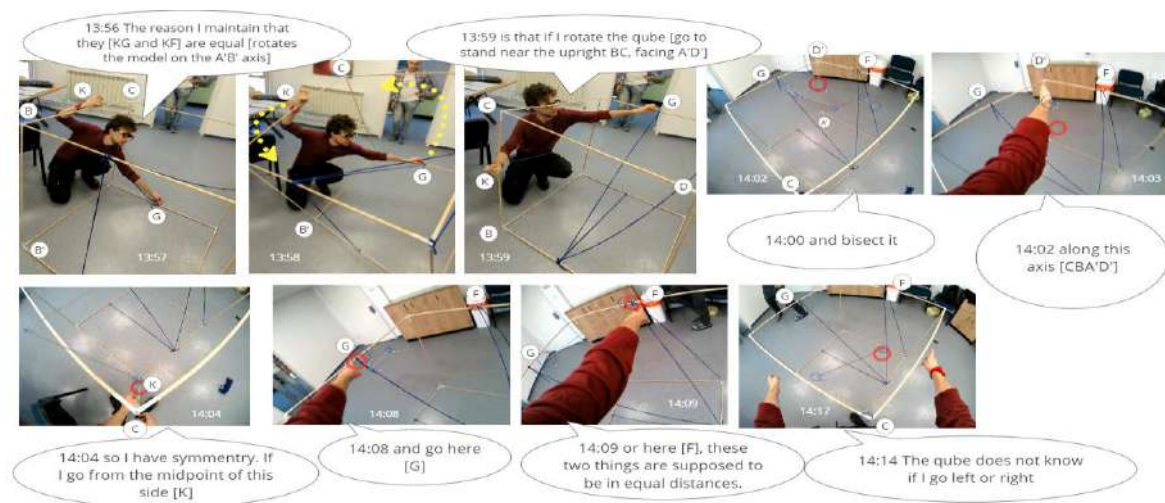


Figure 3: Embodied argumentation for proving that $KG=KF$

Analysis. We maintain that Edi's actions and speech constitute a case of embodied argumentation because a major part of the argument is based on the relative position of his body to the (rotated) cube that revealed the symmetry along the $BCA'D'$ plan. His speech indicated that he thought of equal segments as manifested in the same distance of his own physical movement from K to the two points ("if I go"). Next, we will show that Edi considered this argument to be convincing for a friend and a skeptic.

Data. At 14:55, before addressing our repeated question about convincing himself, a friend, and a skeptic, Edi briefly described another way of proving that $KG=KF$. He said that he could add CG and CF as auxiliary constructions and show that KCG and KCF triangles are congruent since they have: "the same angles and a mutual side (15:16-19)." However, from these two arguments he developed, he chose to use the embodied argument to convince a friend that $KG=KF$ (15:20):

...I was convinced even before I grabbed this 3D model. Friend [stands near $A'D'$ pointing at C with a bisecting gesture similar to Figure 3, 14:03], I could have told him: 'Look with one eye at this side and at that side [gestures with the hand to left and right sections], and you'll see that these two things are equal.'

Edi then re-explained the embodied argument in more detail and later wrote it as a response to a skeptic in the answers form:

The cube can be sliced along the plane $BCA'D'$ to create two symmetric sections where the segments KF and KG pass through symmetric positions. It is also possible to overlap the triangles KCF and KCG. The 2D drawing is misleading.

Analysis. We see that Edi primarily based his argument for a skeptic on the embodied argument but also mentioned the possibility of triangles' congruence. However, note that the CKG and CKF congruence is not straightforward. To show they are congruent, one should first show that $CG=CF$ (e.g., by showing that $\triangle DCF \cong \triangle C'CG$ by SAS). In the simulated recall, when we asked Edi about the incompleteness of the congruence argument, he said:

I could prove it [$\triangle KCF \cong \triangle KCG$], but when I have symmetry, I am very happy to use it... as a shortcut... This line of thought [the congruence] exists, but in my opinion, the proof with symmetry is equally rigorous and can also convince the skeptical.

Evidently, Edi did not thoroughly consider how to prove the congruence because having written the embodied argument, he did not think it was necessary. For him, the embodied argument was rigorous and convincing enough even to a skeptic.

DISCUSSION

In this study, we investigated the roles of embodied argumentation in mathematical proving, particularly its capacity to convince oneself, a friend, and a skeptic (Mason et al., 2010). In the first episode, Edi convinced only himself by an embodied argument rooted in projection-based reasoning. While effective for personal conviction, it required further steps and formalization to persuade others, resembling the

transformation of Proofs Without Words into rigorous proofs through gap-filling activities (Marco et al., 2022). In contrast, the second episode presented a symmetry-based embodied argument that Edi deemed convincing for a friend and even a skeptic. He considered this argument “equally rigorous” to one involving auxiliary constructions and triangle congruence, illustrating how embodied approaches can supplement traditional formal methods in establishing conviction.

The findings reveal differences in the arguments developed for various audiences. Personal conviction often relies on intuitive, tactile experiences, while convincing others requires clear articulation and structured reasoning (Mason et al., 2010). The symmetry-based argument, for example, leveraged the affordances of the 3D model to communicate ideas visually and effectively. This aligns with Stylianides’ (2017) perspective that proof is inherently a social process, requiring sensitivity to the expectations of diverse audiences. However, relying on embodied argumentation raises pedagogical challenges. While it can scaffold understanding and bridge intuition with formal reasoning, there is a risk of over-reliance on intuitive methods without cultivating an *inner skeptic* (Mason et al., 2010). Encouraging students to evaluate their embodied arguments critically is essential to ensuring that embodied argumentation complements and leads to, rather than replaces rigorous reasoning.

The unique affordances of the 3D model—supporting tactile exploration, symmetry, and transformations—highlight its value for geometry curricula. Utilizing these affordances in advanced topics such as symmetry, projections, and rotations can provide a foundation for formalization, enabling students to transition from embodied to symbolic reasoning. This suggests that embodied argumentation is not merely a tool for beginners but also benefits more proficient learners by enriching their multi-modal reasoning processes (see Livingston, 2006).

This study’s limitations include focusing on a single, mathematically proficient participant. However, the fact that even such a knowledgeable participant used and benefited from embodied arguments underscores their relevance beyond novice learners. Future research should explore how students with varying proficiency levels engage with embodied argumentation and how collaborative settings and varying environments shape its development. Additionally, exploring how teachers can effectively support and encourage embodied argumentation in classroom contexts presents a promising research avenue. Design-based research that examines pedagogical principles for structuring the physical environment and crafting effective teacher interventions could provide valuable insights into the interplay between body-artifact interactions and mathematical learning. Despite its limited scope, this study underscores the significant potential of embodied argumentation to enrich mathematical proving by integrating intuitive and formal reasoning, positioning it as a compelling focus for educational innovation.

Acknowledgments

This research was supported by the Israel Science Foundation (Grant No. 1323/23).

References

- Abrahamson, D., & Sánchez-García, R. (2016). Learning is moving in new ways: The ecological dynamics of mathematics education. *Journal of the Learning Sciences*, 25(2), 203–239. <https://doi.org/10.1080/10508406.2016.1143370>
- Bakker, A. (2018). *Design research in education: A practical guide for early career researchers*. Routledge. <https://doi.org/10.4324/9780203701010>
- Herbst, P., Fujita, T., Halverscheid, S., & Weiss, M. (2017). *The teaching and learning of geometry in secondary schools*. Routledge.
- Livingston, E. (2006). The context of proving. *Social Studies of Science*, 36(1), 39–68.
- Mason, J., Burton, L., & Stacey, K. (2010). *Thinking Mathematically (2nd ed.)*. Pearson Education.
- Marco, N., Palatnik, A., & Schwarz, B. B. (2022). When less is more? Investigating gap-filling in proofs without words activities. *Educational Studies in Mathematics*, 111(2), 271–297. <https://doi.org/10.1007/s10649-022-10159-z>
- Marco, N., & Shvarts, A. (Accepted). Proving in geometry as an embodied process. *Journal for Research in Mathematics Education*.
- Nathan, M. J., & Martinez, C. V. (2015). Gesture as model enactment: the role of gesture in mental model construction and inference making when learning from text. *Learning: Research and Practice*, 1(1), 4–37. <https://doi.org/10.1080/23735082.2015.1006758>
- Ng, O. L., Shi, L., & Ting, F. (2020). Exploring differences in primary students' geometry learning outcomes in two technology-enhanced environments: dynamic geometry and 3D printing. *International Journal of STEM Education*, 7, 1–13.
- Palatnik, A., Abrahamson, D., Baccaglini-Frank, A., Ng, O., Shvarts, A., & Swidan, O. (2023). Theory and practice of designing embodied mathematics learning. In M. Ayalon, B. Koichu, R. Leikin, L. Rubel, & M. Tabach (Eds.), *Proceedings of the 46th PME Conference*, (Vol. 1, pp. 159–189). PME.
- Palatnik, A., Abrahamson, D., & Lomos C. (2025). I see what you're thinking: Embodied collaborative argumentation. In G. Bolondi, F. Ferretti, & C. Spagnolo (Eds.), *Proc. 14th Congress of the European Society for Research in Mathematics Education*. ERME.
- Streeck, J., & Mehus, S. (2005). *Microethnography: The Study of Practices*. *Handbook of Language and Social Interaction*, 11237, 381–404.
- Stylianides, G., Stylianides, A., & Weber, K. (2017). Research on the teaching and learning of proof: Taking stock and moving forward. In J. Cai, (Ed.), *Compendium for research in mathematics education* (pp. 237–266). National Council of Teachers of Mathematics.
- Walkington, C., Chelule, G., Woods, D., & Nathan, M. J. (2019). Collaborative gesture as a case of extended mathematical cognition. *The Journal of Mathematical Behavior*, 55, 100683. <https://doi.org/10.1016/j.jmathb.2018.12.002>

I BEG TO DIFFER: THE GERM CELL OF MATHEMATICAL DEBATE

Amedeo Matteucci, Ilaria Veronesi, Francesco Saverio Tortoriello

University of Salerno, Italy

In this theoretical study, we examine how some basic philosophical concepts of Cultural-Historical Activity Theory can aid in understanding the essential peculiarities of debate. Drawing from the history of Activity Theory—that has its roots in Hegel and Marx, before Vygotsky and Leont’ev—we show how it leads us to investigate what constitutes the germ cell and unit of analysis of the debate itself. Having identified a candidate, we argue that it can function as a new lens for designing and analysing debate-based learning activities—e.g. in the collective evaluation of AI-generated outputs, as in our previously proposed activity model. We maintain that this approach also opens fresh avenues for future research, urging educators to harness debate as a transformative social and pedagogical tool.

INTRODUCTION

The diffuse availability of Large Language Model Chatbots—such as ChatGPT, Gemini, Claude, just to name a few—has revolutionised the current educational landscape. Learners have at their disposal a tool that “consistently outperforms the median student grades across various exam scenarios” (Udias et al., 2024) and, since they are in an educational system that typically rewards performance over effort (Chan, 2024), they have no incentive to refrain from leveraging AI for assistance, even fully delegating their work to it.

Nonetheless, as Ted Chiang wittily wrote, “using ChatGPT to complete assignments is like bringing a forklift into the weight room; you will never improve your cognitive fitness that way” (2024, p. 1). The point is that the mechanical reproduction of patterns is not equivalent to understanding, as Searle anticipated with the Chinese Room argument (Searle, 1984) long before the development of advanced AI tools. This is particularly true in mathematics, a discipline whose mastery cannot be reduced to memorising definitions and possessing purely procedural skills. On the other hand, it has been known since antiquity that explanatory dialogue is the principal tool for highlighting the non-mechanical nature of human knowledge (cf. Plato, 1987). For these reasons, our research group already investigated the possibility of using dialogical practices in the classroom to problematise AI use in mathematics education (Matteucci et al., 2024a; Matteucci et al., 2024b), including metamathematical debate (Matteucci & Tortoriello, 2023).

In this respect, i.e. dialogically assessing the validity of AI output, teachers cannot claim a particular expertise above students, since the technology behind AI chatbots

is relatively new and continuously advancing. Even those with formal education or more experience might lack an inherent advantage because expertise in this field quickly becomes outdated. Keeping up with the rapid pace of change requires constant adaptation, not just prior knowledge. Thus, it was natural for us to frame our inquiry in the framework of cultural-historical activity theory (Roth et al., 2012), in particular in Luis Radford's incarnation of it in which, in the context of joint labour, "the focus shifts from how students receive knowledge [...] to how teachers and students co-produce themselves as subjects" (Radford, 2021, p. 24).

But how does the specificity of the activity theory enable us to utilise the debate in classroom as a disciplinary methodology in mathematics? And how could it help us to improve our understanding of the difference between debate and other dialogical practices in mathematics education, e.g. Bartolini Bussi's mathematical discussions (Bartolini Bussi, 1996)?

LITERATURE REVIEW AND THEORETICAL FRAMEWORK

Before addressing directly the issues described in the introduction, in this section we will briefly review the main references in the literature about activity theory, or more properly Cultural-Historical Activity Theory (henceforth CHAT). Then, we will review some of the most important—from our perspective—dialogical practices in mathematics education, namely: mathematical discussion and debate. Finally, we will explain what our stance is in the theoretical framework section, illustrating the theoretical constructs we have chosen to address the core issue of this paper and concluding with the formulation of specific research questions.

Activity Theory

In the *Encyclopedia of mathematics education* we can find the following definition:

Activity theory is the result of an attempt to construct a psychology that draws on and concretely implements epistemological principles of materialist dialectics as K. Marx presented them [...]. Like Marx's *Das Kapital*, activity theory is intended to explain change, learning, and development as an immanent feature of a system rather than in terms of externally produced cause-effect relations. (Roth, 2020, p. 20)

How this attempt at construction should be implemented is not agreed upon in the academic world. Thus, activity theory still lacks a uniquely shared identity and remains more of an umbrella term (Levant et al., 2024) for a series of methodological approaches that are sometimes different, yet originate from a common history—a history that will now be briefly summarised.

Activity theory is rooted in the philosophy of Hegel. It would be impossible to summarise even the basics of Hegel's dialectical philosophy in a few words, however it is important to underline that, for our purposes, the main concepts lie in Hegel's Logic, which is not to be understood as formal logic—or, worse, a substitute for formal logic—but as an ontology that also explains the reflective structure of thought describing concepts. First, basic regularities are abstracted from raw perception,

forming a collection of measures. Then, theories emerge to interpret this data, undergo conflict, and evolve through deeper investigation. Finally, a fundamental abstract concept emerges, encapsulating the phenomenon as a whole. From this “cell,” the phenomenon is reconstructed into an interconnected “organism” by exploring inherent contradictions and their interactions (Blunden, 2021). While Marx diverged from Hegel by rejecting his idealism and emphasising material conditions as the basis of societal change, he retained Hegel’s dialectical method, adapting it to analyse historical and social processes, as he explicitly stated in the introduction of the first volume of *Das Kapital* (Marx, 1987). In this method we can identify two phases: an analytical one where the phenomenon of interest is deconstructed to identify its “germ cell” or fundamental unit, revealing the core contradictions that drive its dynamics; then a synthetic phase where study progresses from the germ cell to reconstruct the phenomenon in its full concreteness. For example, Marx considered the commodity as the germ cell (he used the term *Zellenform*) of capitalist production, encapsulating the contradictions between use-value and exchange-value.

Vygotsky sought to use Marx’s method in psychology and was the first to use the term “unit of analysis” (Vygotsky, 2012) to refer to the smallest recognisable entity that preserves the key features of the whole phenomenon of study and at the same time represents the embryonic concept of it, as germ cell. Traditionally, Vygotsky is considered to be part of the so-called first generation of activity theorists, although “he never proposed or conceptualised activity as a basic unit of analysis” (Engeström, 2024). The first to do this was Leont’ev, one of his collaborators, starting what was later called the second generation. The legacy of these Soviet psychologists was continued in the West by Cole and, later, Engeström, who, drawing also on the results of other followers of Vygotsky, such as Il’enkov and Davydov, expanded the scope of activity theory into what is now often referred to as the third generation. This generation focused on addressing complex, multi-activity systems and the interplay between their components. Engeström introduced the activity system model, which represents activity as a dynamic structure comprising subject, object, tools, community, rules, and division of labour, all mediated by contradictions (Levant et al., 2024). In education, particularly mathematics education, activity theory has been instrumental in analysing how learning unfolds within broader sociocultural contexts. (Roth et al., 2012).

Discussion and debate in mathematics education

Bartolini Bussi defines mathematical discussion in the following way:

Mathematical Discussion is a polyphony of articulated voices on a mathematical object (e.g. a concept, a problem, a procedure, a structure, an idea or a belief about mathematics), that is one of the motives of the teaching-learning activity. The term voice is used after Wertsch (1991), following Bakhtin, to mean a form of speaking and thinking, which represents the perspective of an individual, i.e. his/her conceptual horizon, his/her intention and his/her view of the world. (1996, p. 16).

Debate is another kind of dialogical interaction. According to the book *Argumentation and debate*:

Debate is the process of inquiry and advocacy, a way of arriving at a reasoned judgment on a proposition. Individuals may use debate to reach a decision in their own minds; alternatively, individuals or groups may use it to bring others around to their way of thinking. Debate provides reasoned arguments for and against a proposition. It requires two competitive sides engaging in a bipolar clash of support for and against that proposition (Freeley & Steinberg, 2013, p. 6)

Literature on debate in mathematics is scarce: aside from a few non-academic efforts focused on best practices rather than scientific discourse, one of the few academic contributions to the topic is our previous study (Matteucci & Tortoriello, 2023). Although both debate and mathematical discussion are dialogical interactions, there are some formal differences between them. First, debate is a *structured* dialogical exchange, where participants cannot arbitrarily decide when or if they will contribute. Second, the perspectives presented by debaters are not necessarily reflective of their personal views, they can be assigned by the teacher to allow the exchange to move beyond a personal level to a more genuinely dialectical one. In doing so the teacher acts not only as referee or mediator of the debate but as a guarantor, not of the result, but of the intentionality of the collective activity (Matteucci et al., in press).

Our framework and questions

As researchers in the field of education, we embrace Radford's *communitarian ethics* that "seeks to offer reference points towards the renewal and imagination of new forms of relating to others and working with others" (2021, p. 206) in a perspective of joint labour between students and teachers. This perspective is naturally related to that of the so-called *participant observer* (Blunden, 2023), a stance that involves actively engaging in the context being studied, contributing to its activities, and simultaneously maintaining a critical, reflective perspective. These characteristics of activity theory are due to its Hegelian roots. Therefore, since we are looking at how CHAT can help us understand the specificity of debate as a methodology and also to understand the substantial distinctions—not only the formal ones, that were already outlined above—with other dialogical practices such as mathematical discussion, we make the following statement by Andy Blunden, our own:

Activity Theory is not "what activity theorists think." It is what follows with necessity from the foundations laid by Hegel, Marx and Vygotsky and a commitment to social justice and human emancipation (2023, p. XI).

Thus, if we want to follow the method laid by Marx, and used by Vygotsky and his intellectual continuators, to investigate debate in the context of CHAT, we have to focalise on the concepts of germ cell and unity of analysis. Therefore, we pose the following questions: what is the germ cell of debate as an activity, what is its unit of analysis? How is it possible to capture the specificity of debate with respect to other dialogical practices in mathematics education?

THE GERM CELL AND UNIT OF ANALYSIS OF DEBATE

According to Yrjö Engeström (Engeström, 2024) the germ cell should be: 1) the smallest and simplest initial unit of a totality, serving as its foundational building block; 2) a carrier of the foundational contradiction of the whole, as well as the seed for transcending that contradiction; 3) ubiquitous, and often taken for granted; 4) a perspective opener for multiple applications, extensions, and future developments, enabling diverse interpretations and innovations as the system evolves; 5) action-based and actionable, existing and being enacted in practice before it is explicitly discovered.

We maintain that for debate, the germ cell is: *substantive difference*. Here is our argument for it. At its most minimal level, debate arises from a difference of stance or perspective that compels participants to engage, challenge, or persuade each other. This difference—be it a question, a claim, a proposition, or a dispute—already implies a tension or contradiction: if one says “yes”, the other says “no” to a proposition or, even better, “not quite”. That difference is the tiny seed carrying the entire dynamic of debate within it. With respect to the five characteristics listed above: 1) it is small and simple: a disagreement, a clash of perspectives on a proposition is as simple as it gets; 2) it carries the foundational contradiction that is the driving motor of the practice; 3) it is omnipresent and often overlooked, in the classroom one constantly encounters micro-level disagreements or disputes on how to tackle a question, subtle differences of interpretation and the like; 4) it opens up a perspective for multiple applications since once there is a difference that matters, we have the germ of an entire debate; 5) it is action-based, since people are already arguing or negotiating differences in practice long before they theorise debate as a formal structure. Thus, difference is the germ cell, because from that small contradiction, you can generate the entire complex phenomenon of debate. It is worth noting, however, that not all differences *per se* qualify. A trivial or inconsequential—i.e. whose outcome or resolution does not matter to participants—difference may not spur us to test arguments, offer proof, or attempt to convince an audience: this is why we added the adjective *substantive* to the term *difference*.

We have already said that germ cell and unit of analysis are basically the same thing in CHAT, although in the first case the emphasis is on the developmental aspect, in the latter case it is on the analytical aspect. Therefore we can also conclude that substantive difference also acts as the unit of analysis for the debate. But it is known that some authors identified other constructs as unit of analysis for discursive phenomena. For example “Bakhtin uses the utterance (speaking turn) as the unit for discourse analysis, and Robert Brandom takes the proposition as his unit of analysis” (Blunden, 2021, p. 52). So, why cannot we take these units, instead of substantive difference, to analyse debate? The answer lies in the distinction (Vygotsky, 2012) between analysis by elements—which dissects a phenomenon into its smallest, isolated components, losing its functional integrity—and analysis by units—which

identifies the smallest indivisible entity that retains the essential characteristics and systemic properties of the phenomenon as a whole. Vygotsky uses the famous example of the molecule of water— H_2O —that, while being the smallest unit of water that still retains the characteristics of water, if analysed by its constitutive elements presents us with hydrogen and oxygen, which, individually, do not exhibit the properties of the object under examination. Likewise, if we analyse debate using as unit the utterance or the proposition, we lose sight of the bipolar clash that characterises it—see the above quote from Freeley and Steinberg—whose unity and essence can be reestablished only by combining at least two (contrasting) utterances or propositions—just, as in the case of water, hydrogen and oxygen.

The previous comment on the specificity of the unit of analysis of debate sheds also light on what sets apart debate and mathematical discussion. Referring back to Bartolini Bussi's earlier definition, it clearly appears that the fundamental unit of mathematical discussion is the voice—much like the voices in a polyphonic musical composition (Bartolini Bussi, 2009)—explicitly understood in the sense of Bakhtin's utterance. Thus, the key element of *discussion* is the stance of the individual, in its independence and subjectivity; while the key element of *debate* is the clash of positions, in its relational surpassing of the individuals' personal opinions.

CONCLUSIONS AND IMPLICATIONS FOR THE RESEARCH

In this paper we posited *substantive difference* as the germ cell and unit of analysis of debate. As germ cell, substantive difference helps in devising debatable motions for classroom debates—e.g. by considering differences of opinion on the best method for solving an exercise or on the explanatory power of one way of proving a theorem compared to another. Thus, avoiding the risk of having decontextualised debate motions imposed by the teacher where there is nothing actually to debate. As unit of analysis, substantive difference helps in evaluating the debate already conducted looking for its actual clash points—i.e. focal issues where opposing arguments directly collide, prompting debaters to engage most deeply in defending or refuting a contention—and not for just the initial arguments or, worse, the preconceived notions of the teacher or the audience.

In particular, going back to the activity model we proposed to evaluate AI-generated output (Matteucci et al., 2024b), having identified the germ cell, provides valuable insights for the teacher on how to manage the transition from the classroom discussion phase about the AI's responses to motion drafting. Similarly, knowing what the unit of analysis is gives us guidance on what to look for during the debriefing phase to be later examined in the final discussion.

We also stated that the specificity of debate over mathematical discussion is that its unit of analysis is the said substantive difference, and not the utterance/voice. This conclusion is useful not only for differentiating between the two approaches but also for determining their potential complementarity, after having already established their

compatibility (Matteucci et al., in press). Possibly, also expanding on the interiorization of crucial aspects of theoretical knowledge in a supplementary way to the “voices and echoes game” (Boero et al., 1998)—but this is a subject for possible future research.

Finally, from the previously outlined perspective of the participant observer guided by communitarian ethics, we maintain that the results detailed here are also important because they help teachers organise classroom activities in which students—debating on how mathematical objects historically and socially emerged and evolved—can indirectly address the science-technology contradiction. A contradiction which, while not primary as the labour-capital one that permeates all present-day society, is one of the few leverageable contradictions in mathematics classroom that might bring both social awareness and the social change, for which this awareness is a prerequisite.

References

- Bartolini Bussi, M. G. (1996). Mathematical discussion and perspective drawing in primary school: To Giovanni Prodi on occasion of his 70th birthday. *Educational Studies in Mathematics*, 31(1), 11–41.
- Bartolini Bussi, M. G. (2009). In search for theories: Polyphony, polysemy and semiotic mediation in the mathematics classroom. In M. Tzekaki, M. Kaldrimidou, & H. Sakonidis (Eds.), *Proc. 33rd Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 2, pp. 121–128). PME.
- Blunden, A. (2021). *Hegel, Marx and Vygotsky: Essays on social philosophy*. Brill.
- Blunden, A. (2023). *Activity theory: A critical overview*. Brill.
- Boero, P., Pedemonte, B., Robotti, E., & Chiappini, G. (1998). The “voices and echoes game” and the interiorization of crucial aspects of theoretical knowledge in a Vygotskian perspective: Ongoing research. In A. Olivier & K. Newstead (Eds.), *Proc. 22nd Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 2, pp. 120–127). PME.
- Chan, W. (2024). AI chatbots and liberal education. *AI & Society*, 39(5), 2607–2609.
- Chiang, T. (2024, August 31). Why AI isn’t going to make art. *The New Yorker*. <https://www.newyorker.com/culture/the-weekend-essay/why-ai-isnt-going-to-make-art>
- Engeström, Y. (2024). *Concept formation in the wild*. Cambridge University Press.
- Freeley, A. J., & Steinberg, D. L. (2013). *Argumentation and debate*. Cengage Learning.
- Levant, A., Bakhurst D, Blunden A, Brunila M, Camillo J, Chaiklin S, Cole M, De Smet B, Engeström Y, Hasse C, Rodrigues AM, Mattos C, Nardi B, Oittinen V, Rantavuori J, Sannino A, Stetsenko A, Yamazumi K (2024). *Activity theory: An introduction*. ibidem-Verlag.
- Marx, K. (1987). *Das Kapital. Kritik der Politischen Ökonomie. Erster Band, Hamburg 1872* (6). Walter de Gruyter GmbH. (Original work published 1872)

- Matteucci, A., & Tortoriello, F. S. (2023). Mathematics and debate: a dialectical approach. *Thirteenth Congress of the European Society for Research in Mathematics Education (CERME13)*, Article 18.
- Matteucci, A., Veronesi, I., & Tortoriello, F. S. (2024a). A new agora: The role of AI-shaped debate in transforming mathematics education. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.), *Proc. 47th Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 1, p. 186). PME.
- Matteucci, A., Tortoriello, F. S., Veronesi, I. (2024b). Dialectical chatbots: AI-driven debate in mathematical education. In Faggiano E., Clark-Wilson A., Tabach M., Weigand H.-G. (Eds.), *Proceedings of the 17th ERME Topic Conference Mathematics Education in the Digital Age - MEDA4*, University of Bari Aldo Moro
- Matteucci, A., Tortoriello, F. S., Veronesi, I. (in press). Debating the machine: Mediations and truth. *Fourteenth Congress of the European Society for Research in Mathematics Education (CERME14)*.
- Plato. (1987). *Theaetetus ; Sophist* (Fowler, H.N., Trans.). Harvard University Press. (Original work published ca. 370 B.C.E.)
- Radford, L. (2021). *The theory of objectification: A Vygotskian perspective on knowing and becoming in mathematics teaching and learning*. Brill.
- Roth, W.-M., Radford, L., & LaCroix, L. (2012). Working with cultural-historical activity theory. *Forum: Qualitative Social Research, Volume 13, No. 2*, Art. 23.
- Roth, W.-M. (2020). Activity theory in mathematics education. In S. Lerman (Ed.), *Encyclopedia of mathematics education*, (p. 20–23).
- Searle, J. R. (1984). *Minds, brains and science*. Harvard University Press.
- Udias, A., Alonso-Ayuso, A., Alfaro, C., Algar, M. J., Cuesta, M., Fernández-Isabel, A., Gómez, J., Lancho, C., Cano, E. L., de Diego, I. M., & others. (2024). ChatGPT's performance in university admissions tests in mathematics. *International Electronic Journal of Mathematics Education*, 19(4).
- Vygotsky, L. S. (2012). *Thought and language, revised and expanded edition*. MIT Press. (Original work published 1934)
- Wertsch, J. V. (1991). *Voices of the mind: Sociocultural approach to mediated action*. Harvard University Press.

THE ROLE OF DOMAIN-GENERAL COGNITIVE SKILLS IN EARLY MATHEMATICAL DEVELOPMENT: A LONGITUDINAL STUDY OF 4- AND 5-YEAR-OLD CHILDREN

Méndez Jurado, María Agustina and Pearson, Rufina

CIPP-UCA (Centro de Investigaciones de Psicología y Psicopedagogía- Universidad Católica Argentina)

This study investigates the relationship between domain-general cognitive abilities and mathematical achievement in 4- and 5-year-old children. Using a longitudinal design, assessments were conducted at two time points to assess cognitive and mathematical skills, including visual working memory, oral comprehension, and visual-spatial abilities. Regression models showed that oral comprehension was the primary predictor of counting, whereas both oral comprehension and visual-spatial skills had a significant influence on conventional knowledge and arithmetical operations. Visual working memory did not develop as a relevant predictor. These results highlight the developmental progression of mathematical skills and underscore the role of domain-general cognitive skills in supporting early numeracy.

INTRODUCTION

Numerical concepts are acquired in children from an early age and form the basis of further education (Aragón et al., 2022; Hornung et al., 2014). Dehaene and Cohen (1995) proposed the Triple Code Theory, which organizes numerical representation into three modules: magnitude (non-symbolic estimation), verbal (linguistic representation for counting and basic arithmetic), and visual (Arabic numeral processing for complex operations). These systems develop separately between the ages of two and four, but they are progressively incorporated into formal education to improve numerical accuracy and efficiency (Aragón et al., 2022).

Mathematical skills include both domain-specific skills and domain-general abilities. Domain-specific early numerical abilities include a variety of tasks such as counting, Arabic numeral recognition, number comparison, number ordering, and calculation (Verschaffel et al., 2023). These skills are critical for engaging with numerical concepts and represent the building blocks of mathematical development. On the other hand, domain-general skills support learning across different abilities, such as memory, attention, and reasoning (Geary, 2017). Working memory, spatial abilities, fluid intelligence and general language skills act as important resources that support mathematical performance in all contexts (Hornung et al., 2014; Verschaffel et al., 2023).

The aim of this study is to examine the relationship between domain-general cognitive skills and mathematical skills in 4- and 5-year-old children. A longitudinal approach is used to examine how early cognitive and mathematical skills develop over time and

contribute to later mathematical achievement. The following research questions are asked: to what extent do domain-general cognitive skills relate to mathematical skills in preschool-aged children (4–5 years); how much early cognitive and mathematical skills assessed at age 4 predict mathematical performance at age 5; and which domain-general cognitive skills are most influential in the development of mathematical abilities in early childhood.

METHOD

Research Design

As part of a longitudinal PhD research project including from preschool to 1st grade, this study focuses on the early stages of data collection to assess the predictive relationship between domain-general cognitive skills and mathematical skills. Future phases of the project will incorporate additional abilities, such as verbal working memory, quantitative reasoning, and subsequent mathematical performance, to expand the understanding of these complex developmental processes. This study uses a longitudinal research design that tracks the development of mathematical competencies over a two-year period.

Participants

The study involved 133 children ages 4 and 5 (55.6% female), attending private kindergarten schools in Buenos Aires, Argentina. Assessments were conducted at two different times points. Children were assessed first at 4 years of age and then, the same cohort at 5 years of age. The average age of the children was 59.5 months ($SD = 3.39$) in Year 1, comprised of 4-year-old children, and 67.6 months ($SD = 3.41$) in Year 2, comprised of 5-year-old-children. Participants were selected through convenience sampling, ensuring diversity in socioeconomic and educational backgrounds. Parental consent was obtained for all participants, and ethical protocols were strictly followed throughout the study.

Data Collection Instrument

The digital version of the JEL MAT K screening (Pearson & Pearson, 2014) was used as the primary instrument for assessing mathematical skills. This screening tool is an adaptation of the Number Sense Brief (NSB) Screener, developed by Jordan et al. (2008), a validated and widely used measure of early numerical abilities. The NSB is a condensed version of the Number Knowledge Test (NKT) and has demonstrated robust predictive validity, with correlations of .63 between preschool assessments and third-grade math achievement. The Digital JEL MAT K Screening includes 12 subtests covering i.e. counting skills, conventional numerical knowledge, arithmetical operations and symbolic and non-symbolic representations. This tool is particularly suitable for longitudinal research because it focuses on foundational numerical concepts, which are critical to early mathematical development.

Procedure

To ensure optimal performance, children were familiarized with the examiners. Each child was individually assessed during school hours in two 30-minute sessions, with breaks provided whenever the child showed signs of fatigue. In the first year of the study, when participants were 4 years old, assessments focused on both domain-general and domain-specific cognitive skills. Visual working memory was assessed with the Visual Span Subtest from WPPSI-IV (Wechsler, 2014). Domain-specific mathematical skills were assessed using the Digital JEL MAT-K Screening.

In the following year, the same cohort of children was reassessed to examine the relationship between cognitive and mathematical skills. Oral language skills were assessed using the Woodcock-Muñoz III test (Schrack et al., 2005), which provides a comprehensive measure of verbal expression and language comprehension. In addition, visuospatial abilities were assessed with the Block Design subtest from WPPSI-IV (Wechsler, 2014), which examines spatial reasoning and the ability to analyse and synthesize abstract visual stimuli. The tests were conducted and scored by experts in Child Neuropsychology and a PhD student in Educational Psychology.

Data analysis

Descriptive statistics were calculated to summarize the data, followed by a Shapiro-Wilk test (González-Estrada et al., 2022) to assess the normality of the sampling distribution. The results showed that the sample had a normal distribution for most variables as evidenced by the Shapiro-Wilk test ($p > .05$). Regression analyses, conducted in Jamovi (The Jamovi Project, 2024), identified domain-general cognitive skills that predict performance in counting, conventional knowledge, arithmetical operations, and overall mathematical achievement in children aged 4 and 5. This approach enabled exploration of the predictive power of cognitive abilities on mathematical skills, provided understanding of the specific contributions of different cognitive domains to early numeracy development.

RESULTS

Descriptive analyses were performed to summarise the key variables, followed by an Exploratory Factor Analysis (EFA) to examine how the variables clustered. The analysis identified four distinct factors. The first factor, named *Counting*, included the subtests Concrete Counting (.851), Progressive Oral Counting (.689), Backward Counting (.500), and Associating Symbolic and Non-Symbolic Magnitudes (.333). The second factor, termed *Conventional Knowledge*, comprised subtests such as Number Recognition (.584), Number Writing (.502), and Symbolic Magnitude Comparison (.389). The third factor included Verbal Calculations (.469), Non-Verbal Calculations (.796), and Subitizing (.626), while the fourth factor was represented exclusively by Problem Solving (.995). Based on Pearson and Pearson's (2014) framework for early mathematical skills, the third and fourth factors were combined into a single

subcomponent named *Arithmetical Operations*. The factor loadings for each subtest and their corresponding factors are summarized in Table 1 below.

Table 1

Factor Loadings for Mathematical Subtests in the Exploratory Factor Analysis

	Counting	Conventional Knowledge	Arithmetical Operations
Concrete Counting	.851		
Progressive Oral Counting	.689		
Backward Counting	.500		
Number Recognition		.584	
Associating Symbolic and Non-Symbolic Magnitudes	.333		
Verbal Calculations			.469
Problem Solving			.995
Number Writing		.502	
Non-Verbal Calculations			.796
Symbolic Magnitude Comparison		.389	
Subitizing			.626

Mathematical achievement and its subcomponents were assessed across both years. The standard deviations indicate greater variability in performance during Year 1 compared to Year 2. Notably, performance in arithmetic operations showed significant improvement in Year 2, reflecting overall growth in mathematical skills. The descriptive statistics are presented in Table 2.

Table 2

Domain-specific Mathematical Skills Descriptive Statistics

		Mathematical Achievement	Counting	Conventional Knowledge	Arithmetical Operations
Year 1	Mean	42.4	8.10	42.5	5.67
	Median	42.2	9	36.0	5.29
	Standard Deviation	17.7	2.72	30.0	3.30

Year 2	Minimum	5.10	1	4	0.06
	Maximum	80.3	11	158	13.3
	Mean	57.0	9.28	39.1	8.58
	Median	55.3	10	37	8.26
	Standard Deviation	16.6	2.23	12.5	3.54
	Minimum	8.2	2	5	0.18
	Maximum	83.3	19	58	15.3

Note: N=133

The Shapiro-Wilk test was used to evaluate whether the data followed a normal distribution. The test indicated that the data follows a normal distribution ($W = 0.989$, $p = 0.401$), justifying the use of parametric statistical methods (González-Estrada et al., 2022).

To examine the predictive relationship between domain-general cognitive skills and total mathematical performance, a series of regression analyses were conducted.

Relationship Between Total Mathematical Achievement and Domain-general Cognitive Skills

A linear regression analysis was conducted to examine the predictive relationship between cognitive abilities and mathematical achievement in 4-year-old children. The model revealed that oral comprehension ($\beta=1.684$, $p<.001$) and visual-spatial skills ($\beta=0.982$, $p=.004$) were significant predictors of mathematical achievement, explaining 26.5% of the variance ($R^2 = 0.265$, $R^2_{adjusted} = 0.254$). Similarly, oral comprehension ($\beta=1.697$, $p<.001$) and visual-spatial skills ($\beta=0.982$, $p=.004$) were significant predictors in mathematical achievement in 5-year-old-children, while visual working memory did not contribute significantly ($\beta=-0.120$, $p=.795$). The final model explained 26.6% of the variance in mathematical achievement ($R^2 = 0.266$, $R^2_{adjusted} = 0.248$).

Relationship Between the Subcomponents and Domain-general Cognitive Skills

A separate regression analysis was performed to explore the relationship between cognitive abilities and the subcomponents. Regarding counting in 4-year-old-children, the results indicated that oral comprehension ($\beta = 0.275$, $p<.001$) was a significant predictor of counting scores. However, visual-spatial skills did not reach significance in this model ($\beta=0.066$, $p=.239$). The model that includes visual-spatial skills and oral comprehension explained 20.2% of the variance ($R^2=0.202$, $R^2_{adjusted}=0.189$). In the case of 5-year-old-children, the model explained 14.3% of the variance ($R^2=0.143$,

$R^2_{adjusted}=0.123$). Oral comprehension ($\beta=0.1787$, $p<.001$) was the only significant predictor.

Regression analysis examining conventional knowledge in aged 4 and 5 showed that both oral comprehension (year 1: $\beta=1.90$, $p=.002$; year 2: $\beta=1.074$, $p<.001$) and visual-spatial skills (year 1: $\beta=2.53$, $p<.001$; year 2: $\beta=0.747$, $p=.005$) significantly predicted performance. The model in the first year accounted for 19.7% of the variance ($R^2=0.191$, $R^2_{adjusted}=0.1791$). Whereas the regression model for conventional knowledge in 5-year-old-children accounted for 20.6% of the variance ($R^2=0.206$, $R^2_{adjusted}=0.188$).

Finally, a regression analysis was conducted with arithmetical operations as the dependent variable. In the case of 4-year-old-children, oral comprehension ($\beta=0.358$, $p<.001$) and visual-spatial skills ($\beta=0.213$, $p=.001$) were both significant predictors. This model explained 30.5% of the variance ($R^2=0.305$, $R^2_{adjusted}=0.294$). Similarly, in 5-year-old-children both oral comprehension ($\beta=0.443$, $p<.001$) and visual-spatial skills ($\beta=0.1591$, $p=.018$) emerged as significant predictors. This model explained 35.1% of the variance ($R^2=0.351$, $R^2_{adjusted}=0.336$).

CONCLUSIONS

This study highlights the role of domain-general cognitive skills, particularly oral comprehension and visual-spatial skills, in predicting mathematical achievement in 4- and 5-year-old children. Regression analyses revealed that these abilities were consistent predictors across both mathematical achievement and specific subcomponents such as conventional knowledge and arithmetical operations. In contrast, visual working memory did not contribute significantly, suggesting its limited role in early mathematical development.

The findings also show that oral comprehension, as a verbal measure, was the primary predictor of counting, while both oral comprehension and visual-spatial skills significantly predicted conventional knowledge and arithmetical operations. The variance increased with more complex tasks, such as arithmetical operations, which highlights the integration of linguistic and spatial abilities as mathematical tasks became more sophisticated.

This research demonstrates the developmental progression of mathematical skills, with improvements observed across all measures from age 4 to 5. These results emphasize the important role of domain-general cognitive abilities during early childhood to support the acquisition of foundational mathematical skills.

References

Aragón, E., Canto-López, C., Aguilar, M., Menacho, I., & Navarro, J. (2022). Longitudinal study of symbolic and non-symbolic magnitude processing and its relationship with mathematical achievement. *Revista de Psicodidáctica*, 28(1), 44-50.

- Dehaene, D. y Cohen, L. (1995). Towards an anatomical and functional model of number processing. En Butterworth, B., *Mathematical Cognition*, 83-120. Psychology Press.
- Geary, D., Nicholas, A., Li, Y., & Sun, J. (2017). Developmental change in the influence of domain-general abilities and domain-specific knowledge on mathematics achievement: An eight-year longitudinal study. *Journal of Educational Psychology*, 109(5), 680-693.
- González-Estrada, E., Villaseñor, J., & Acosta-Pech, R. (2022). Shapiro-Wilk test for multivariate skew-normality. *Computational Statistics*, 37.
- Hornung, C., Schiltz, C., Brunner, M., & Martin, R. (2014). Predicting first-grade mathematics achievement: The contributions of domain-general cognitive abilities, nonverbal number sense, and early number competence. *Frontiers in psychology*, 5(272), 1-18.
- Pearson, R., Pearson, J., (2014). *Manual de Programa de Entrenamiento en Habilidades Matemáticas Jel-Matk*. Bibliográfika.
- Schrank, F. A., McGrew, K. S., Ruef, M. L., Alvarado, C. G., Muñoz-Sandoval, A. F., & Woodcock, R. W. (2005). *Batería III Woodcock-Muñoz: Overview and technical supplement (Assessment Service Bulletin No. 1)*. Itasca, IL: Riverside Publishing.
- The jamovi project (2024). *jamovi*. (Version 2.6) [Computer Software]. Retrieved from <https://www.jamovi.org>.
- Verschaffel, L., De Smedt, B., Luwel, K., Onghena, P., Torbeyns, J., & Van Dooren, W. (2023). Development and stimulation of early core mathematical competencies in young children: Results from the Leuven Wis & Co project. En *Mathematical Challenges For All* (R. Leikin, pp. 25-50). Springer.
- Wechsler, D. (2012). *Wechsler Preschool and Primary Scale of Intelligence-Fourth Edition (WPPSI-IV)*; Pearson: San Antonio, TX, USA.

MATHEMATICAL SELF-CONCEPT AND EXPERIENCE OF COMPETENCE WHILE MODELLING WITH AND WITHOUT EXPERIMENTS

Marielena Menzel¹, Timo Kosiol², Stefanie Rach², Sebastian Geisler¹

¹University of Potsdam, ²Otto-von-Guericke University Magdeburg

The current study investigates the potential to foster students' mathematical self-concept when working on modelling tasks by combining them with experiments. In addition, we identify the mechanisms that underlay this potential by considering basic needs and the reciprocal relation between self-concept and achievement. N=627 upper-secondary students participated in an experimental study with control-group design. We analyse data from questionnaires measuring self-concept using a mediation model and including experience of competence as a mediator. The results support a reciprocal relationship between self-concept and experience of competence. However, we can't identify an effect of combining modelling tasks and experiments on self-concept compared to solely working on a modelling task.

INTRODUCTION

Applying mathematics to solve real-world problems is relevant for modern society (Niss, 1994). Mathematical modelling is a process where a real-world problem is solved by translating it into mathematics and finding solutions for the real-world problem. As schools should prepare students to participate in modern society, teaching mathematical modelling has become part of the curricula (OECD, 2023). However, students' motivation regarding mathematical modelling does not reflect the relevance of that area. Studies show, for example, that students are not confident they can solve modelling tasks (Krawitz & Schukajlow, 2018). Thus, students' self-concept should be fostered. One approach to foster mathematical self-concept is to combine modelling tasks with experiments (Beumann, 2016). This combination provides a hands-on activity contributing real data and also allows experiencing the origin of the modeled data and, therefore, offering a better understanding. As self-concept emerges from mastery experiences in the past, students need to experience competence while working on modelling tasks (Bong & Skaalvik, 2003). In the current study, firstly, we investigate the impact of combining modelling tasks with experiments and of students' learning prerequisites on their mathematical self-concept. Secondly, we explore the mechanism of these effects, assuming that experience of competence when working on modelling tasks has an impact.

THEORETICAL BACKGROUND

Students' self-concept and experience of competence

A person's self-concept of ability (SCA) is conceptualized as a "more general stable estimate of one's global but domain specific ability. [...] task specific success expectation derived from the individual's domain specific SCAs" (Eccles & Wigfield, 2024, p. 51). The self-concept of ability is therefore conceptualized as someone's perception of their capability to perform in the questioned area, which in the current study is mathematics. Bong and Skaalvik (2003) argue that different factors influence a person's self-concept. One factor is a person's mastery experiences, which express experiences of succeeding at a task. The student's experience of competence (known from basic needs, Ryan & Deci, 2000) can conceptualize this mastery experience. One approach to foster students' self-concept is, therefore, to provide experiences of competence. Another factor is actual achievement. A recent review conducted by Breit et al. (2025) found support for a relation between self-concept and achievement in mathematics (see also Marsh et al., 2022). Thus, we can state that both experience of competence and achievement may influence someone's self-concept. Grades or knowledge test scores can capture the achievement. A test score gives direct information about the achievement in the area of interest and the grades via teachers' judgement are an indirect representation of students' achievement, where we cannot control which facets have been taken into account. Thus, in the current study, we focus on knowledge test scores as they contribute more reliable information about the actual cognitive capacity than prior grades given by teachers.

Mathematical modelling with experiments and students' motivation

When solving a real-world problem with mathematical modelling, several steps are needed (Leiss et al., 2024). At first, the real problem is described in a task, which has to be understood, simplified, and structured before it is translated into a mathematical model, where it can be solved mathematically. The mathematical results have to be interpreted and retranslated into the real world. The whole process needs validation. It has been criticized that modelling tasks in schoolbooks often use smoothed data, which fit the intended model perfectly, thus distorting the impression of mathematical application (Engel, 2018). Nevertheless, prior research shows that students do not feel capable of solving modelling tasks (Krawitz & Schukajlow, 2018). This result coincides with the observation that working on modelling tasks is challenging for students (Hankeln, 2019). One approach to lowering the challenging character of modelling tasks is to add experiments to modelling tasks. When combining modelling tasks with experiments, students conduct a hands-on experiment themselves, where they collect data about the given problem. The self-collected data is then used for choosing a mathematical model. Thus, experiments represent the real-world situation and contribute data that has to be modeled in a subsequent task (Halverscheid, 2008). Previous research indicates that combining modelling tasks with experiments can foster

students' motivation (Ganter, 2013; Geisler & Rach, 2023) and especially students' self-concept and their experience of competence (Beumann, 2016).

However, these findings do not allow conclusions as to which of the facets of experimenting the effects are attributable: the hands-on activity of conducting an experiment or the working with real data. Furthermore, it is unclear whether the findings are replicable when applying a control-group design. Thus, we designed the current intervention study with students working on modelling tasks with and without experiments and with and without real data to investigate each facet's influence.

RESEARCH INTEREST – THE CURRENT STUDY

According to prior research, students' mathematical self-concept needs fostering when working on modelling tasks. As self-concept relates to the experience of competence, we want to promote students' self-concept by providing learning situations of modelling tasks with experiments, which could lead to a higher experience of competence. In contrast to typically used smoothed data in schoolbooks (Engel, 2018), which fit the intended model ideally, using data from self-conducted experiments provides two differences: the hands-on activity of conducting an experiment and working with real data. By working with three condition groups in an intervention, we control both effects that experimenting brings into the task. In the current study, we want to investigate firstly the impact of the intervention on the self-concept when controlling the learning prerequisites and secondly the mechanism of this effect through experience of competence. This leads to the research questions:

1. In what way does conducting an experiment and working with real data (self-collected and given) vs. given smoothed data as well as students' learning prerequisites (pre-self-concept and prior knowledge) affect the self-concept after the intervention?

Hypothesis 1: (a) Conducting an experiment will positively affect the post-self-concept (Beumann, 2016). For working with real-data and given smoothed-data, we have no such assumption. (b) The pre-self-concept and the prior knowledge will have a significant positive effect on the post-self-concept.

2. To what extent does the experience of competence mediate the effects on post-self-concept (direct and indirect effects)?

Hypothesis 2: The relations between working with different types of data (experimental data, real data, or smoothed data) or students' learning prerequisites (pre-self-concept and prior knowledge) and post-self-concept can be mediated by the experience of competence.

METHODOLOGY

Design and Sample

The current study is part of the project *Experiments to foster modelling competences and motivation in mathematics* (Ex2MoMa), where $N = 627$ students (51% female,

$M_{age} = 16$ years) from 27 upper secondary school classes participated. In three consecutive mathematics lessons the classes worked on three modelling tasks with different contexts, all solvable using functions – one task in each lesson. Every class was randomly assigned to one of three conditions. The conditions varied only in the type of data the students had to use for the modelling tasks: (i) self-collected experimental data, (ii) given real data, and (iii) given smoothed data (see Figure 1).

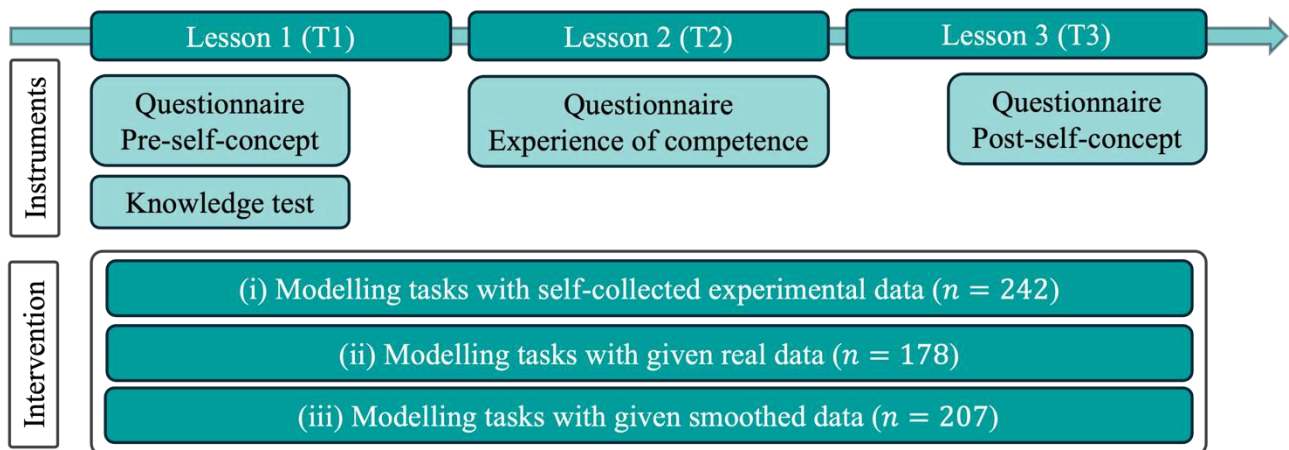


Figure 1: Design of the current study

At the beginning of the first lesson and at the end of the third lesson, all students filled out a questionnaire about their self-concept regarding mathematics and at the end of the second lesson, they answered a questionnaire about their experience of competence concerning the modelling task they worked on right before. In addition, we conducted a test to capture students' prior content knowledge about functions at T1.

Instruments and Analysis Strategy

The questionnaires at T1 and T3 are identical and consist of one approved scale of four items about self-concept (item-example: "I am good at mathematics.", Arens et al., 2011). The reliabilities for both measurement points are good (T1, T3: $\alpha > .93$). The questionnaire T2 consists of one approved scale of four items about the experience of competence with good reliability (T2: $\alpha = .91$; item-example: "While working on the task, I felt competent.", Willems, 2011). All answers for the three questionnaires were given on a 6-point Likert scale (1 = not at all true, 6 = totally true). The prior knowledge about functions was captured with a knowledge test with 7 dichotomous-coded items (single-choice and open answer mixed, see Figure 2).

An exponential function f is given in the value table. State the corresponding functional equation for f .

x	0	1	2	3
$f(x)$	1	3	9	27

$f(x) = \underline{\hspace{2cm}}$

Figure 2: Item-example for the content knowledge test regarding functions

The answers have been analyzed using item-response-theory methods. The WLE person parameter was used as control variable, with a higher value representing higher achievement ($WLE_{rel} = .62$).

We applied regression-based analyses to answer the research questions. For the first research question, we used students' *post-self-concept* as dependent variable and *working with different conditions*, *pre-self-concept*, and *prior knowledge* as independent variables. To investigate the effect of the different conditions, we used two dummy variables: “conducting an experiment” (modelling with experiments = 1, modelling with given real data = 0, modelling with smoothed data = 0) and “real data” (modelling with experiments = 1, modelling with given real data = 1, modelling with smoothed data = 0). By doing so, we can differentiate the unique influence of the hands-on experimenting activity and separate it from the influence of working with real data. To answer the second research question, we use mediation analysis, with *experience of competence* as mediator of the relation between the predictors (identified in the previous analysis) and *post-self-concept*. Not all students participated in all three lessons causing missing data. The analyses were calculated in Mplus (Version 8), and we used the FIML algorithm for missing data for both dependent variables. As we allocated the students by class to the three conditions, we adjusted the standard errors with *type = complex* in Mplus.

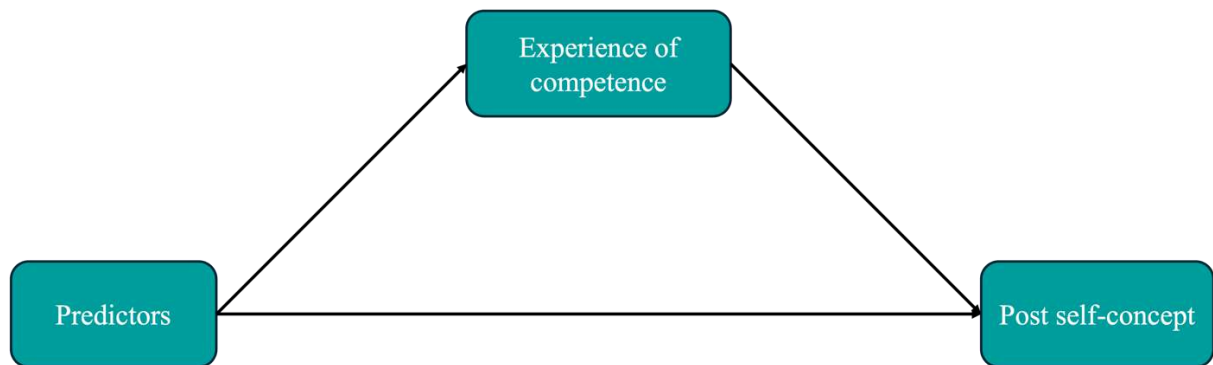


Figure 2: Model for mediation analysis

RESULTS

By answering the first research question, we sought to identify predictors of self-concept. We found the following results: For both dummy variables “conducting an experiment” and “real data”, the standardized path coefficients were not significant. Thus, we cannot identify any effect of either working with any particular type of data nor of a hands-on activity on students' self-concept. In contrast to that, *pre-self-concept* ($N = 448$, $\beta = .878$, $p < .001$) and *prior knowledge* ($N = 448$, $\beta = .077$, $p < .01$) predict *post-self-concept*. The *pre-self-concept* has a strong effect on the *post-self-concept*, whereas the *prior knowledge* has a rather weak effect. That means a higher *pre-self-concept* and higher *prior knowledge* coincides with a higher *post-self-concept*. Therefore, we cannot confirm hypothesis 1a but find support for hypothesis 1b.

Regarding the second research question, we wanted to identify the mechanism behind the predictors' effects on the *post-self-concept* – by considering the mediator *experience of competence*. As we cannot give evidence for the impact of the different conditions on the *post-self-concept*, we do not consider them for the mediation analysis. We find significant effects for the standardized path coefficients from the predictors *pre-self-concept* and *prior knowledge* on the mediator *experience of competence* and from *experience of competence* on the *post-self-concept*, all relations with small effect sizes (see Figure 3).

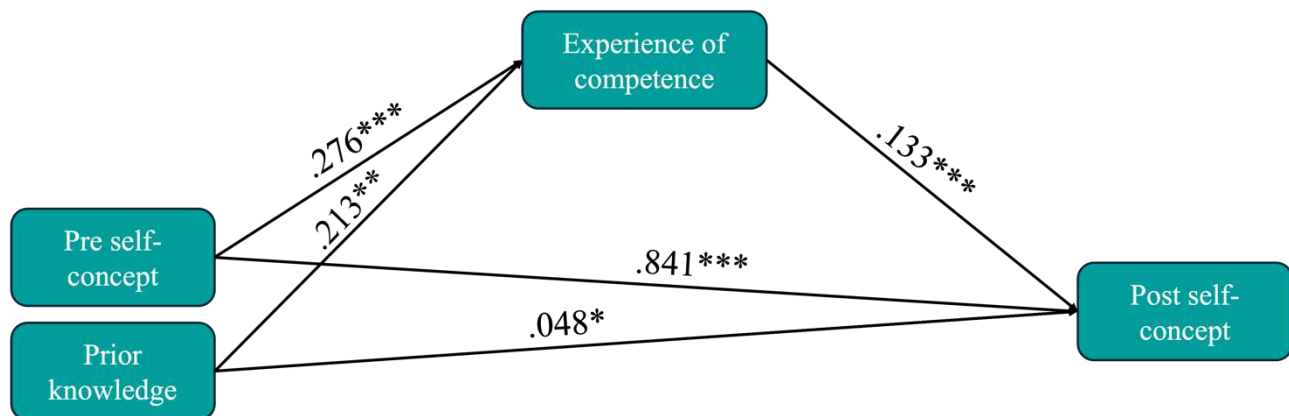


Figure 3: Standardized path coefficients and direct effects, $N = 516$, $R^2 = .835^{***}$,
 $*p < .05$, $**p < .01$, $***p < .001$

We used bootstrap confidence intervals (based on 5000 bootstrap samples) to test the indirect effects. The effect is significant if zero does not lie within the bootstrap confidence interval. We find significant positive indirect effects with small effect sizes for the *pre-self-concept* ($\beta = .037$, 95% CI [0.020; 0.062]) and for *prior knowledge* ($\beta = .028$, 95% CI [0.014; 0.049]) via the mediator *experience of competence* on the *post-self-concept*. Besides this mediation by *experience of competence*, *pre-self-concept* directly affects *post-self-concept* with a large effect size and *prior knowledge* affects the *post-self-concept* with a small effect size (see direct path in Figure 3). As the direct effect of the *individual characteristics* on the *post-self-concept* is significant, the influences of the predictors cannot only be explained via *experience of competence*. Thus, we find partial support for hypothesis 2.

DISCUSSION

The current study focuses on fostering students' self-concept by combining modelling tasks with experiments. We found that mathematical self-concept is a slightly stable construct, and we cannot identify different effects of the conditions in the intervention (conducting an experiment, working with real data and working with smoothed data) on students' mathematical self-concept (in contrast to the findings of Beumann, 2016). The divergent findings may be explained by the fact that the intervention of the current study with three mathematics lessons within three weeks is relatively short and that longer-term interventions may have the potential to substantially influence the self-concept. According to the results of Marsh et al. (2022), we support the link between

mathematical self-concept and achievement by predicting self-concept by prior content knowledge, controlling for prior self-concept. Moreover, in the current study we wanted to identify the mechanism underlying the effects on post-self-concept. Here, we investigated the relationship between students' self-concept and their experience of competence while also taking prior knowledge as predictor into account. Therefore, we connect prior research by Bong and Skaalvik, (2003) and Marsh et al. (2022), which – taken together – theoretically propose a reciprocal influence of these constructs: students' self-concept and their prior knowledge influence how competent the students feel when working on mathematical tasks which is indicated by the experience of competence. These experiences influence their latter mathematical self-concept. In the current study, we could identify such a mechanism for modelling tasks. The assumed relation of the prior knowledge via the experience of competence on the self-concept seems to be a good way of capturing that link. Summing up, the learning prerequisites that we cover with prior knowledge and pre-self-concept are more critical for the post-self-concept than the concrete design of the groups in the intervention.

Due to the intervention's short duration, the current study's findings regarding the impact of modelling with and without experiments are limited. The results of the current study are also limited to modelling tasks with functions, and it remains open whether they are transferable to other types of mathematical tasks or other mathematical domains like geometry.

Future studies could investigate whether the motivation-fostering potential of experimentation for self-concept would be supported if the experimentation was aligned with corresponding feedback and reflection of the students' achievement. Another interesting perspective of future research would be to analyze long-term developments of student's self-concept regarding modelling tasks when these tasks are regularly applied in mathematics lessons. Despite the intervention did not affect the student's self-concept, experimenting combined with modelling still could positively influence other motivational factors and could, therefore, be considered to be applied in the classroom but has to be reflected by the teacher, depending on the desired outcome. Moreover, modelling with experiments could have a fostering impact on cognitive student variables like modelling competencies.

Funding: This study is funded by the German federal ministry of education and research under grant No. 16MF1028B.

References

- Arens, K., Trautwein, U., & Hasselhorn, M. (2011). Erfassung des Selbstkonzepts im mittleren Kindesalter. [Capturing middle-aged children self-concept] *Zeitschrift für pädagogische Psychologie*, 25, 131-144.
- Beumann, S. (2016). *Versuch's doch mal: Eine empirische Untersuchung zur Förderung von Motivation und Interesse durch mathematische Schülerexperimente*. [Just try it: an empirical study concerning the promotion of motivation and interest with mathematical experiments] (Dissertation). Ruhr-Universität Bochum.

- Bong, M., & Skaalvik, E. M. (2003). Academic Self-Concept and Self-Efficacy: How Different Are They Really? *Educational Psychology Review*, 15(1), 1–40.
- Breit, M., Schneider, M., & Preckel, F. (2025). Mathematics achievement and learner characteristics: A systematic review of meta-analyses. *Learning and Individual Differences*, 118, 102621.
- Eccles, J., & Wigfield, A. (2024). The development, testing, and refinement of Eccles, Wigfield, and Colleagues' Situated Expectancy-Value Model of Achievement, Performance and Choice. *Educational Psychological Review*, 36, 51
- Engel, J. (2018). *Anwendungsorientierte Mathematik: Von Daten zur Funktion* [Application-oriented mathematics: From data to function.] Springer Berlin Heidelberg.
- Ganter, S. (2013). *Experimentieren - ein Weg zum Funktionalen Denken: Empirische Untersuchung zur Wirkung von Schülerexperimenten*. [Experimentation - a path to functional thinking: Empirical study on the effect of student experiments.] Kovač.
- Geisler, S., & Rach, S. (2023). Students' interest when combining modelling and experimentation—Is it worth the effort? In M. Ayalon, B. Kouichu, R. Leikin, L. Rubel, & M. Tabak (Eds.), *Proc. of the 46th Conference of the International Group for the Psychology of Mathematics Education Vol. 2*, 363–370).
- Halverscheid, S. (2008). Building a local conceptual framework for epistemic actions in a modelling environment with experiments. *ZDM*, 40(2), 225–234.
- Hankeln, C. (2019). *Mathematisches Modellieren mit dynamischer Geometrie-Software: Ergebnisse einer Interventionsstudie*. [Mathematical modelling with dynamic geometry software: results of an intervention study.] Springer Fachmedien Wiesbaden.
- Krawitz, J., & Schukajlow, S. (2018). Do students value modelling problems, and are they confident they can solve such problems? Value and self-efficacy for modelling, word, and intra-mathematical problems. *ZDM*, 50(1–2), 143–157.
- Leiss, D., Ehmke, T., & Heine, L. (2024). Reality-based tasks for competency-based education: The need for an integrated analysis of subject-specific, linguistic, and contextual task features. *Learning and Individual Differences*, 114, 102518.
- Marsh, H. W., Pekrun, R., & Lüdtke, O. (2022). Directional Ordering of Self-Concept, School Grades, and Standardized Tests Over Five Years: New Tripartite Models Juxtaposing Within- and Between-Person Perspectives. *Educ Psychol Rev* 34(4), 2697–2744.
- Niss, M. (1994). Mathematics in Society. In R. Biehler, R. W. Scholz, R. Strässer, & B. Winkelmann (Eds.), *Didactics of Mathematics as a Scientific Discipline*. Kluwer Academic Publishers.
- OECD. (2023). *PISA 2022 Assessment and Analytical Framework*. OECD.
- Ryan, R. M., & Deci, E. L. (2000). Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *American Psychologist*, 55(1), 68–78.
- Willems, A. S. (2011). *Bedingungen des situationalen Interesses im Mathematikunterricht* [Conditions of situational interest in mathematics teaching.] Waxmann.

FOSTERING PRESERVICE TEACHERS' MWS- INTERPRETATIVE KNOWLEDGE THROUGH EPISTEMOLOGICAL TAXICAB PERTURBATIONS

Annamaria Miranda¹, Gonzalo Espinoza², Romina Menares³, Paula Verdugo⁴

¹Università di Salerno, Fisciano, Italy amiranda@unisa.it, ²Universidad Alberto Hurtado, Santiago, Chile gespinoza@uahurtado.cl, ³Universidad de Valparaíso, Chile romina.menares@uv.cl, ⁴Universidad de Talca, Linares, Chile pauverdugo@utalca.cl

Klein's discontinuity is a significant issue in teacher education research, focusing on the gap between university-acquired mathematical knowledge and the necessary knowledge for teaching. This paper investigates prospective teachers' interpretation of geometric tasks, involving the transition from Euclidean geometry to Taxicab geometry, using Interpretative Knowledge (IK) and Mathematical Working Spaces (MWS) to reduce this gap. Moreover, the MWS-IK relationship is introduced and used.

INTRODUCTION AND CONCEPTUAL BACKGROUND

Klein's discontinuity refers to the gap between the mathematical knowledge acquired in university courses and the mathematical knowledge a teacher should have to teach (MKT) (Grenier-Boley & Robert, 2024). Teacher knowledge research builds on Shulman's work (1986). Noteworthy is the Carrillo's MTSK model (2008), which organises mathematical teacher specialised knowledge (MTSK) as an organic whole, considering epistemological aspects and adaptation to educational contexts. It addresses teachers' evidence and needs, allowing understanding of their teaching approach. Ribeiro et al. (2013) introduced interpretative knowledge (IK) as a part of mathematical knowledge that helps teachers understand non-standard or error-containing answers, enabling them to comprehend students' mathematical productions. It is a construct referring to the knowledge needed to be able to make sense of students' mathematical productions. The theory of Mathematical Working Space (MWS) defines a mathematical working space as an abstract structure organised to account for mathematical work and enable individuals to solve problems in a specific domain. Espinoza-Vasquez et al. (2024) examine the integration of MWS and MTSK to understand teachers' classroom practice, highlighting the potential of these constructs in addressing practical issues (Henríquez-Rivas & Espinoza-Vásquez, 2018). This paper introduces the MWS-IK construct, combining IK and MWS, by intertwining semiotic representations, instrumental artefacts, and discursive referential, as well as the induced processes and planes, with teachers' IK'. We discuss how prospective teachers interpret peers' processes activated by shifting from Euclidean to Taxicab geometry. We face the issue of promoting prospective teachers' MKT by giving them educational opportunities that allow them to develop interpretative knowledge in a broader way (MWS-IK) stimulated by a geometry domain changing.

THEORETICAL FRAMEWORK

The framework for designing and for analysing is based both on MWS and IK theories as well as on their merging when perturbing the epistemological plane. The design of tasks aims to create conditions for the development of the cognitive processes included in the MWS theory and to favour the development of the interpretative knowledge approach. The geometric transition from Euclidean to Taxicab geometry even naturally stimulates non-standard reasoning and a complete circulation within MWS.

Taxicab geometry

Exploring concepts in Taxicab geometry in the context of geometry can help students better understand concepts in Euclidean geometry (Miranda & Saliceto, 2024). The Euclidean distance between two points is the length of the straight-line segment connecting them, while, to calculate Taxicab distance, the sum of the horizontal and vertical distances between them is used. Euclidean distance (d_E) and Taxicab distance (d_T) between two points $A = (a_1, a_2)$ and $B = (b_1, b_2)$ are defined as:

$$d_E(A, B) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2} \quad d_T(A, B) = |a_1 - b_1| + |a_2 - b_2|$$

MWS theory

MWS theory aims to analyse the mathematical work performed by an individual when solving tasks through a model considering the epistemological and cognitive planes of objects studied within mathematical domains (Kuzniak et al., 2022). On the epistemological plane, there are three components: *representamen*, which refers to the representation of the mathematical object; *artefact*, the materials or symbolic systems that can be used as means of action; and *theoretical referential*, which specifies mathematical definitions and properties. On the cognitive plane, components include *visualisation*, the perception of the mathematical object through different semiotic systems; *construction*, referring to actions triggered by artefacts used and associated techniques; and *proof*, referring to discursive reasoning. MWS also considers relations between components of the epistemological and cognitive planes by means of three geneses. *Semiotic genesis* is based on representation registers, allowing for their identification and processing and conversion between different registers. *Instrumental genesis* enables the operationalisation of artefacts in the construction process. Finally, *discursive genesis* allows for mathematical reasoning based on definitions and properties. In MWS, interactions between geneses and their associated components leads to the activation of three vertical planes: *semiotic-instrumental* [Sem-Ins], when artefacts are used to build under certain conditions or to explore semiotic representations; *instrumental-discursive* [Ins-Dis], involving proofs based on experimentation, exploration, or justifying a construction; and *semiotic-discursive* [Sem-Dis], where the proof is coordinated with the visualisation process.

Interpretative knowledge

The construct of Interpretative Knowledge introduced by Ribeiro et al. (2013) refers to a deep and wide mathematical knowledge that enables teachers to support students in building their mathematical knowledge starting from their own reasoning and productions, without how not standard or incorrect they might be. IK completes the knowledge of typical errors or solution strategies with the knowledge of possible sources for errors and the knowledge of possible uses of errors. IK also includes the ability to develop specific feedback based on the sense given to the students' reasoning; it should allow them to exploit the potential of erroneous or unexpected strategies.

MWS-IK relationship. A first approach

Asenova et al. (2023) introduce the notion of semiotic interpretative knowledge (SIK) and envisage this theoretical tool as a building block of a mathematics teacher's specialised knowledge. They characterise SIK as 'the knowledge needed by teachers in order to interpret students' answers [...] when conceptual knowledge is hindered and, thus, remains hidden, behind difficulties related to patterns of *sign use* and production, including individual creativity in sign use.' In different words, SIK seems to correspond to 'the knowledge required to interpret, make sense of, and explore students' productions, particularly those that are non-standard and based on errors occurring during the semiotic genesis and hindering the visualisation.' Analogously, we refer to *Instrumental Interpretative Knowledge* (IIK) as 'the knowledge required to interpret, make sense of, and explore students' productions, particularly those that are non-standard and based on errors occurring during the instrumental genesis and hindering the construction.' And naturally, in coherence with the theory of MWS, *Discursive Interpretative Knowledge* (DIK) is 'the knowledge required to interpret, make sense of, and explore students' productions, particularly those that are non-standard and based on errors occurring during the discursive genesis and hindering the proving.' The definition of the mathematical working space interpretative knowledge, which we denote with MWS-IK, is outlined. It is declined by SIK, IIK, and DIK and their interconnections as the knowledge required to interpret, make sense of, and explore students' productions, particularly those that are non-standard and based on errors, occurring when a student circulates from the epistemological to the cognitive plane of the MWS model when engaged in a mathematics activity. Interactions between geneses and their associated components lead to the activation of the interpretative knowledge within three *vertical planes*: *semiotic-instrumental* [Sem-Ins]-IK, *instrumental-discursive* [Ins-Dis]-IK, and *semiotic-discursive* [Sem-Dis]-IK.

RQ *What kind of mathematical work is encouraged moving from Euclidean to Taxicab geometry? How do prospective teachers develop MWS-Interpretative knowledge?*

METHODOLOGY

This research was conducted in a post-university preservice teacher training course, lasting 10 hours, attended by 30 graduates, held during the second semester of the academic year 2023–24 in a southern Italy university. They worked in the domain of geometry (Kuzniak & Nechache, 2021), divided into groups, and free to use any resource or tool useful and discuss the solution on vertical non-permanent boards (Table 2). A workshop activity consisting of three main tasks was carried out collectively. In the first task, each group, once recalling some Euclidean concepts, was asked to solve problems concerning the reality, model the problem through the Taxicab distance, define the Taxicab metric, and reason in the new geometry through the construction of examples, counterexamples, and proofs. The second task was based on the interpretation of the work produced by another group. Finally, the third one required designing a path to be addressed to secondary school students.

Method of data analysis

We focus our qualitative content analysis on the identification of MWS-IK signs within the WG1' answers in a guide sheet requiring the interpretation of the WG2' solution.

MWS-IK	Description
Coordination	Corresponds to teacher interpretative knowledge of:
n	
SIK	the semiotic dimension going from the identification of the suitable semiotic representation to the representation to the mathematical object (visualisation).
IIK	the instrumental dimension going from the identification of the suitable artefacts and moving towards the construction triggered by the artefacts and associated usage techniques.
DIK	the discursive dimension that, starting from the use of definitions, properties, or theorems of a theory, enacts a discursive reasoning based on distinct forms of justification, argumentation, or demonstration to construct new products of the theory.
[sem-ins]-IK	the coordination of the manipulation of an artefact (including a technological one) for the construction of a mathematical object with a semiotic representation, or of a semiotic representation with the use of the artefact to explore other different representations.
[sem-dis]-IK	the coordination of the process of visualisation of the objects represented with discursive reasoning to prove a result or define a concept.
[ins-dis]-IK	the coordination of discursive processes coming from empirical proofs with an artefact, or the validation of a construction, or the instrumental processes to validate a theorem.

Table 1: Protocol for MWS-IK analysis.

FINDINGS AND DISCUSSION

WG2 focus on the definition of Taxicab-circle and the second Euclid Theorem.

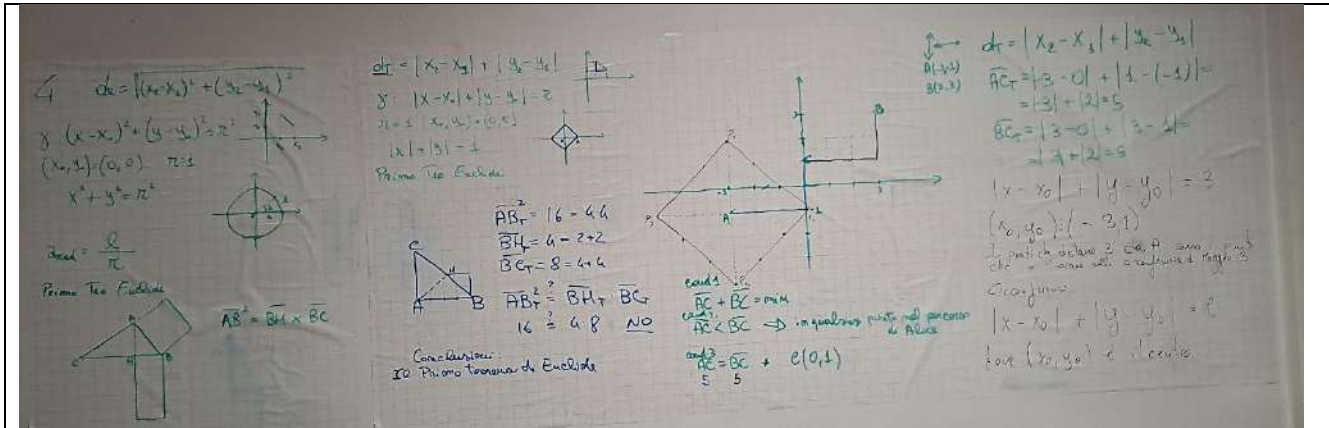


Figure 1: The WG2's solution of task 1 on the board (on the left Euclidean, on the right Taxicab geometry)

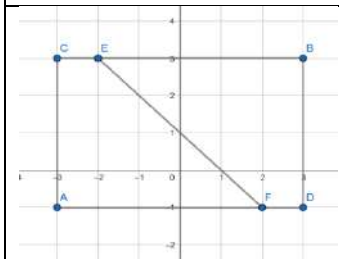


Figure 2: The WG2's graph of the reality problem with GeoGebra (task 1- Episode2)

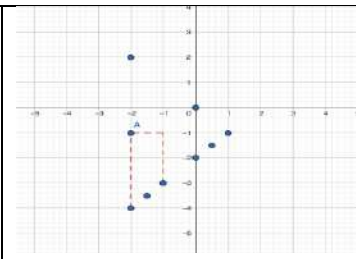


Figure 3: The WG2's Taxicab-circle point by point with GeoGebra (task 1-Episode3)

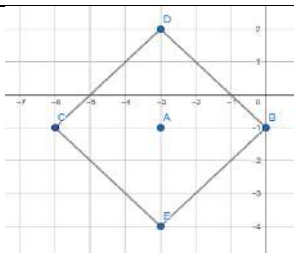




Figure 4: The WG2's Taxicab-circle with GeoGebra (task 1-Episode3)

Table 2: Summary of the resolution of task by WG 2.

WG1, looking at the WG2's board (Figure 1) and reading the WG2'S document, examines the work of the WG2 to interpret its reasoning and provide immediate feedback. We report the comments contained in the interpretation guide sheet (Table 3) in which WG1 is asked to identify how WG2 worked to solve the problem as well as which processes WG2 activated, highlighting the non-standard reasonings or errors and thinking about how to use them to orchestrate a learning collective discussion.

MWS	Episode 1	Episode 2	Episode 3
 ins-dis			The fact that being a circumference depends on the position creates tension with intuition, and the representation using the tool helps to convince oneself and reflect on the definition.
 sem-dis	They give the definition of radiant but do not use it afterwards. This is a discursive obstacle behind the concept of radiant.		When WG2 tries to prove the validity of Euclid's Theorem, since the measurement of the sides in the Taxicab metric depends on the position, this creates difficulties. This obstacle is overcome by considering the algebraic relationship of the Theorem and proving that it does not hold in a particular case.




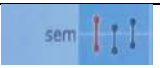
 sem-ins		They represented the problem by points, but to convince themselves, they passed to GeoGebra.	The circumference is represented graphically and analytically, but GeoGebra supports the intuitive break created by the new geometry, allowing for the examination of infinite possible cases. ([sem-ins] – IK)
 dis	The discursive genesis was activated by recalling the definitions of distance, circle, and the statement of Euclid's theorem, but they do not retrace the proof.		The discursive genesis was activated in providing the definition of Taxicab distance, giving the definition of circumference, and showing that Euclid's theorem does not hold. At this point, we identified some critical issues in proving that the theorem does not hold, and we tried to understand what the obstacles depended on. The geometric reasoning in this case is not standard precisely because of how geometry is made (DIK, SIK).
 ins	Epistemological reference to artefacts (GeoGebra) and their role in exploring the definition of a circle, but they do not graph it.	GeoGebra obstacles to identify the rectangle in which the problem finds a solution, probably due to the abandonment of reasoning on the board. (IIK)	They explore with the use of GeoGebra what a circumference means to overcome the obstacles due to the single case treated on the blackboard. (IIK)
 sem	The representations from which we start are the analytical and geometric ones: geometric and analytical representations of Euclidean distance and circle. Synthetical representation of Euclid's theorem		They write and represent the equation of the circle with radius 1 and centre at the origin, then, to prove that Euclid's Theorem is not valid, they encounter difficulties in the graphic-geometric visualisation (SIK), but they are convinced by the numerical example. Reasoning about the proof starting from the graphical case is misleading [sem-ins]-IK. We would have considered the algebraic relationship to construct the counterexample. The choice of semiotic representation could be an obstacle in class. (SIK)

Table 3: Interpretative task elaborated by WG1 on a guide sheet.

We analyse through the lens of MWS-IK the interpretative document to understand how the interpretative knowledge components emerge and are shaped by WG1. Table 4 synthesises the three most significant dimensions of the MWS-IK signs detected.

SIK	IIK	DIK
Episode 1: Euclidean world WG2 avoids Euclid's theorem, giving only a synthetical representation, and this seems to induce a WG1's	WG1 observes that WG2 talks about the digital representation of a circle, but they do not graph it.	WG1 say that the discursive genesis is activated even if WG2 only limits recalling definitions. Valuable observation on the concept of radians, and the possibility of

reflection on the aiding transition to the formal proof.		discussing it in the taxi context to make it understandable.
Episode 2: real life WG1 does not observe that there is no reference to the analytical representation of paths.	WG1 highlight difficulties in reasoning on the board and the need to use GeoGebra.	WG1 does not note that there is no reference made to the definition of taxi distance.
Episode 3: Taxicab world WG1 observes that the choice of semiotic representation could be an obstacle in class. Reasoning only on the graphical case is misleading. WG1 gives feedback saying that as teachers they would have suggested a conversion process to construct the counterexample. Specifically, they realise in this setting the algebraic register works well.	They explore with the use of GeoGebra what a circumference means to overcome the obstacles due to the single case treated on the blackboard	WG1 observes WG2's critical issues in proving that the theorem does not hold and tries to understand what the obstacles depended on. The geometric reasoning in this case is non-standard because of how geometry is structured, and WG1 is trying to interpret the reasons and reflect on how to create a learning occasion.

Table 4: MWS-IK-based analysis.

The study investigates how perturbing the theoretical referential component of Kuzniak's MWS model can enhance prospective teachers' interpretation of peers' work. The MWS is an abstract environment in which a teacher moves by activating useful processes: acquiring the mathematical knowledge he needs to teach, to design, to evaluate. Among them, the interpretative knowledge that a teacher should have to manage non-standard solutions in their discursive, instrumental manifestations stands out. Inspired by both theories, we introduce the notion of MWS-IK connecting MWS and IK, we define the interpretative knowledge connected to components, domains and planes and use it to analyse prospective teachers' productions. It represents a framework for designing tasks useful for prospective teachers' training since they would suggest them practices transforming errors into resources. We design a workshop to be addressed in a prospective teacher training course encouraging their circulation in a mathematical working space as solvers and as peers 'work interpreters simulating that would take place in the classroom. The experience carried out shows that stressing the vertices of the epistemological plane to design the task produces significant solutions in terms MWS-IK. Results show how MWS-IK can foster a broader interpretative knowledge and provide effective feedback in teacher training.

References

- Asenova, M., Del Zozzo, A., & Santi, G. (2023). Unfolding Teachers' Interpretative Knowledge into Semiotic Interpretative Knowledge to Understand and Improve Mathematical Learning in an Inclusive Perspective. *Education Sciences* (1) 65.
- Carrillo, J., Climent, N., Montes, M., Contreras, L. C., Flores-Medrano, E., Escudero-Ávila, D., Vasco, D., Rojas, N., Flores, P., Aguilar-González, A., Ribeiro, M., & Muñoz-Catalán,

- M. C. (2018). The mathematics teacher's specialised knowledge (MTSK) model. *Research in Mathematics Education*, 20(3), 236-253. <https://doi.org/10.4000/adsc.756>
- Espinoza-Vásquez, G., Henríquez-Rivas, C., Climent, N. *et al.* Teaching Thales's theorem: relations between suitable mathematical working spaces and specialised knowledge. *Educational Stud Math* (2024). <https://doi.org/10.1007/s10649-024-10367-9>
- Grenier-Boley, N., & Robert, A. (2024). How can university mathematics overcome Klein's second discontinuity? Specific course design. *ZDM Mathematics Education* 56, 1459-1471. <https://doi.org/10.1007/s11858-024-01604-9>
- Kuzniak, A., & Nechache, A. (2021). On forms of geometric work: A study with pre-service teachers based on the theory of mathematical working spaces. *Educational Studies in Mathematics*, 106, 271-289. <https://doi.org/10.1007/s10649-020-10011-2>
- Kuzniak, A., Montoya Delgadillo, E., & Richard, P. (2022). *Mathematical work in educational context: The perspective of the theory of mathematical working spaces*. Springer. <https://doi.org/10.1007/978-3-030-90850-8>
- Henríquez-Rivas, C., & Espinoza-Vásquez, G. (2018). Relación ETM-MTSK: Conexiones entre la génesis semiótica y el conocimiento de los temas. In E. Montoya, P. Richard, L. Vivier et al. (Eds.), *Espacio de trabajo matemático. Actas ETM6* (pp. 507-512).
- Miranda, A. & Saliceto L. (2024). Back and Forth Between Euclidean Geometry and Taxicab Geometry to Foster Students' Theoretical Thinking in Digital Contexts. *International Journal for Technology in Mathematics Education*, 31(2), 79-87.
- Ribeiro, M., Mellone, M., & Jakobsen, A. (2013). Characterizing prospective teachers' knowledge in/for interpreting students' solutions. In A. Lindmeier & A. Heinze (Eds.), *Proceedings of the 37th PME*, Kiel, Germany.
- Ribeiro, M., Mellone, M., & Jakobsen, A. (2016). Interpreting students' non-standard reasoning: Insights for mathematics teacher education practices. *For the Learning of Mathematics*, 36(2), 8-13.
- Shulman, L.S. Those who understand: Knowledge growth in teaching. *Educ. Res.* 1986, 15, 4-14.

INTER-COMPONENT CHANGES IN MATHEMATICS TEACHING ACTIVITY - FOCUS ON POTENTIAL CORE CHANGE PROCESSES

Victoria Möller

Goethe-University of Frankfurt, Germany

This study investigates teachers' competence in teaching fraction addition and subtraction, focusing on intuitive, algorithmic, and formal components of mathematical activity (Fischbein, 1994). Classroom video analysis reveals that inter-component changes can present potential challenges when transitions are ambiguous. The findings highlight the importance of maintaining mathematical consistency across components to support effective learning. The study underscores the need for targeted teacher training to address these transitions, enhancing teachers' ability to provide more coherent and meaningful learning experiences for students.

INTRODUCTION

Assuming that teachers design mathematics lessons in a way that is as conducive to learning as possible, they can be seen as the central driving force behind mathematics learning in school. To design mathematics lessons, teachers make use of their professional competence, which combines their cognitive abilities and skills as well as their motivational, volitional, and social readiness and skills (Weinert, 2014). One goal of professional teaching-competence models can be described as the promotion of professional teaching action: the performance of teacher action in the concrete implementation in (mathematics) lessons (Kunter et al., 2013).

Challenging learning content places particularly high demands on teachers when designing the teaching–learning situation (Wittwer & Renkl, 2008). Since fractions seem to be particularly challenging learning content (Copur-Gencturk, 2021), this paper investigates in detail the multimodal utterances of the teacher during a lesson on adding and subtracting fractions.

With the components of mathematical teaching activity (Fischbein, 1994), the present study pays particular attention to teachers' statements and their different content design with regard to an intuitive, algorithmic, and formal component. Through statements on all components of mathematical teaching activities, teachers provide a diverse range of learning opportunities that offer different connection potentials. When designing the learning offer, the mathematically consistent presentation and linking of the various components of mathematical teaching activities is particularly important. Previous analyses have shown that, in particular, the transitions between the components are sometimes contradictory and therefore not meaningful for learners (Möller, in press-b). The following section takes a closer look at these potentially challenging teaching situations.

THEORETICAL FRAMEWORK

Components of mathematical activity within mathematical teaching processes

According to Fischbein (1994), mathematics should be understood not only as a formal and deductive procedure but also as a creative human activity. It includes three basic components: the formal, the algorithmic, and the intuitive components of mathematical activity (Fischbein, 1994). In the following, examples, the components of mathematical activity are presented on the basis of the learning content of the teaching unit under consideration: the addition and subtraction of fractions. The formal component involves axioms, definitions, and theorems, for example considering fractions as equivalence classes. The algorithmic component of mathematical activity focuses on (standardized) procedures for solving mathematical problems and thus draws on procedural knowledge (Prediger, 2008). For example, expanding a fraction can be described as a sub-step of the addition–subtraction algorithm. The intuitive component of mathematical activity is necessary to stimulate thought processes and is accepted without further legitimation. This includes, for example, the idea that the further division of a piece of cake does not change its overall size.

Assuming that mathematical activities are inherent to mathematical teaching–learning situations, the components of mathematical activity can be identified here. Since instructional processes are largely initiated and structured by the teacher, it is worth investigating in detail which components of mathematical activity the teacher utters during their mathematics instruction. If the components of mathematical activity are focused on teachers' statements, they are referred to below as components of mathematical teaching activity. Initial analyses suggest that there is a connection between the changes in different components of mathematical teaching activity and potentially challenging teaching content (Möller, in press-a).

Intercomponent changes - a potentially challenging teaching situation

Mathematical consistency refers to the logical non-contradiction of mathematical content in the context of the components of mathematical teaching activities. It can be understood as a central characteristic of the description of meaningful mathematical statements in teaching–learning contexts and can be sharpened in the context of mathematical teaching activities in two complementary dimensions: the intracomponent and intercomponent dimensions (Möller, in press-a; see Figure 1).

In the intracomponent dimension, mathematical consistency ensures that statements within a component of mathematical teaching activity are clearly and comprehensibly related to each other. Consequently, the ideas, strategies, and axioms expressed within the same component of mathematical teaching are also consistent with each other. In addition, mathematical consistency also plays an important role on the inter-component level and thus with regard to the change between different components of mathematical teaching. The aim here is to ensure that the mathematical content expressed between different components of mathematical teaching activity can be linked logically and

argumentatively in a comprehensible and non-contradictory manner. In this way, the argumentative logic of the mathematical content can be presented in a meaningful way for learners through the two dimensions of mathematical consistency (Möller, in press-a).

The changes between the components of mathematical teaching activities are sometimes contradictory in terms of content and therefore appear to be potentially challenging for teachers. Focusing more closely on inter-component changes seems worthwhile in order to further understand these potentially challenging situations.

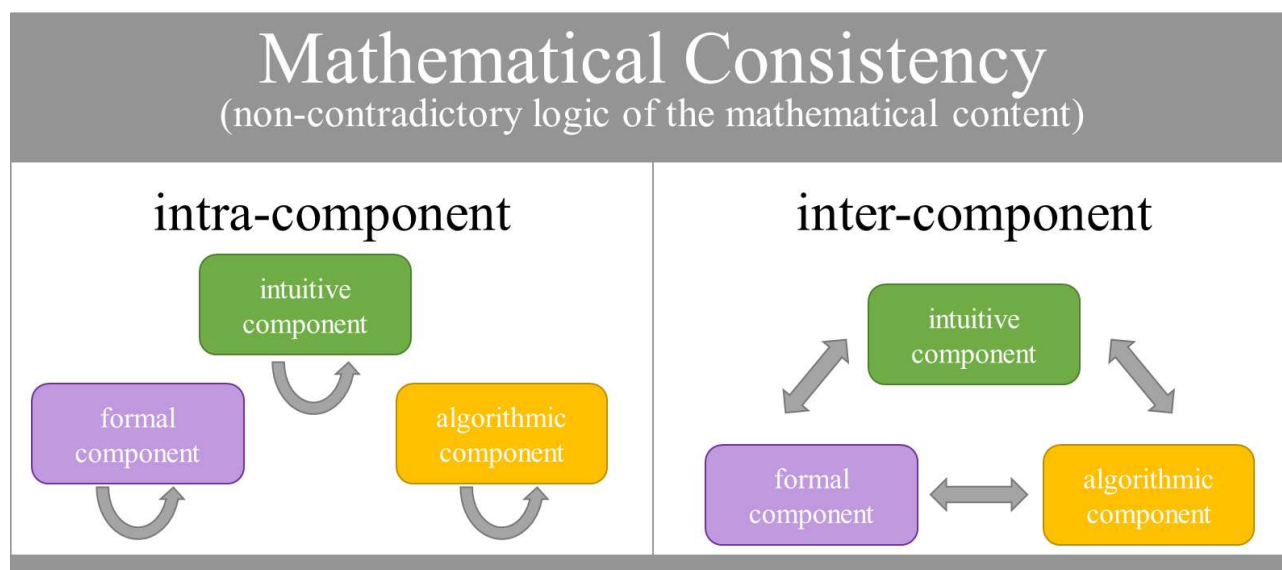


Figure 1: Mathematical Consistency in Two Dimensions

RESEARCH QUESTIONS AND DESIGN

In the presented analysis, the following research questions are investigated: (1) When do inter-component changes in mathematical teaching activity occur in the global temporal–successive course of a lesson? (2) How can the local framework situation of inter-component changes be described when the learning content is named for the first time?

The following section analyzes 80 minutes of a video from a sixth-grade class at an integrated comprehensive school in Germany. During the lesson, the topics of the addition and subtraction of fractions are introduced.

Method of Data Analysis

Two analytical methods are employed to address the research questions. To respond to the initial research question, the data is examined through a structuring content analysis approach, as outlined by Mayring (2014), using a theory-based category system. It is structured deductively along the components of mathematical activity (Fischbein, 1994) and is both created and inductively expanded based on existing analyses (Möller, in press-b). The resulting main categories are labeled *intuitive*, *algorithmic*, *formal*,

ambiguous, and *other*. While *intuitive*, *algorithmic*, and *formal* describe mathematical activities, the main category *ambiguous* includes mathematical content that cannot be assigned to the other (*intuitive*, *algorithmic*, and *formal*) main categories. This is the case, for example, when the teacher makes intuitive and algorithmic statements at the same time. The main category *other* includes organizational or incomprehensible utterances.

Each main category is differentiated into subcategories. For example, the subcategories within the *algorithmic* category describe the individual steps of the addition and subtraction algorithm (*alg-1*: expand, *alg-2*: add/subtract numerator and keep denominator, *alg-3*: reduce the result completely) or a combination of at least two steps of the algorithm (*alg-steps*). Additionally, the decomposition of a fraction into its components outside of an operation context (*alg-fraction*) or an algorithmic expression that cannot be assigned to any of the subcategories described so far (*alg-ambiguous*) also belong to the main algorithmic category.

To answer the second research question, an adaptation of the context analysis (Mayring, 2014; Vogel, 2017) is employed, which enables a reconstruction of the individual mathematical teaching concepts in the observed teaching unit through contrasting analysis. The theoretical foundation for this analysis is provided by Fischbein's (1994) conceptualization of mathematical activity. His theory is concretized for the addition and subtraction of fractions and serves as a reference framework for the analysis. For the present paper, the sequential–thematic insights thus gained (Möller, in press-b) are considered in their naturally occurring chronology.

ANALYSIS INSIGHT AND RESULTS

(1) Global consideration: inter-component changes in mathematical teaching activities

Mayring's (2014) structuring content analysis indicates that the integration of fraction addition and subtraction is predominantly characterized by algorithmic factors. In the observed teaching unit, inter-component changes are marked by a temporary shift away from the algorithmic component of mathematics teaching activities. Regarding inter-component changes in mathematical teaching activities, three sequences can be distinguished in the observed lessons: some, few, and no inter-componential changes (see Figure 2).

At the beginning of the teaching unit, the algorithmic component of mathematics teaching activities is temporarily discontinued on several occasions for the intuitive anchoring of the learning content (00:11:50–00:12:30, 00:22:30–00:22:40). Moreover, these changes also include intervals that were ambiguous with regard to the components of mathematical teaching activities. In some intervals, the teacher clearly utters an individual component of mathematical teaching activity; in other cases, the intervals cannot be assigned to one. It appears that the teacher is providing an

orientation for the localization of learning content across the various components of mathematical teaching activities. As the observed lesson progresses, ambiguous intervals and thus potential inter-component changes are coded. The inter-component changes gradually decrease and become less concrete, the previously introduced learning content stabilizes with regard to the components of mathematical activity. Toward the end of the teaching unit, neither concrete changes between the components of mathematical teaching activities nor ambiguous intervals can be identified. The learning content is expressed throughout the algorithmic component of mathematical teaching activity, which means that the core of this teaching unit, already indicated as algorithmic, is applied. For more detailed analyses, the inter-component changes that occur at the beginning of the lesson appear to be rich; this particularly applies the intervals coded as ambiguous.

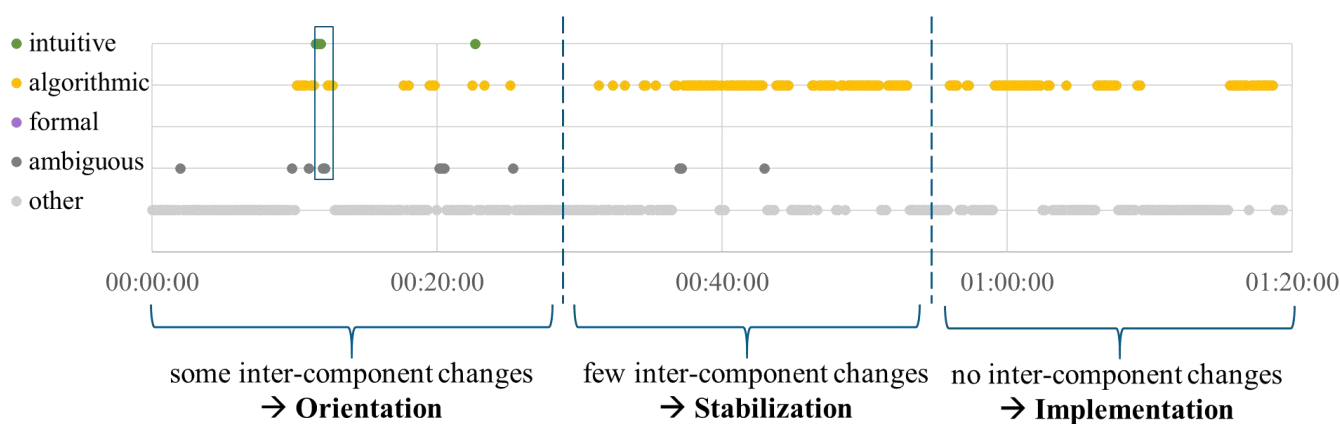


Figure 2: Inter-Component Changes in the Observed Mathematics Lessons

(2) Local consideration: mathematical consistency of inter-component changes in mathematical teaching activities

For a detailed consideration of the inter-component changes, a sequence at the beginning of the lesson is selected as an example of the change between the intuitive, ambiguous, and algorithmic intervals (see rectangle in Figure 2).

For the context analysis, a transcript segment that appears particularly rich with regard to the research question is analyzed first. For the present study, this is a segment coded as ambiguous (00:12:00–00:12:20). This analysis step is followed by the selection and analysis of further passages that enrich the teaching concept.

In the selected segment, the teacher utters through speech, “Add only pieces of the same size” and “I have to divide them into equal pieces” (see gray box in Figure 3), running the edge of her hand over different diameters of an imaginary circle. Compared to the frame of reference, the speech utterance of the sequence can be interpreted both as a formulation of the premise of the addition algorithm (algorithmic) and as an idea of fractions as parts that undergo a (common) equal division (intuitive). Furthermore, the gesture can be seen as a repeated cutting or dividing of a round whole (e.g., a cake).

The segment, therefore, contains aspects of the algorithmic and intuitive components of mathematical teaching. On closer inspection, it is noticeable that the necessity of equal denominators for the algorithmic component of mathematical teaching activities can be confirmed. The teacher, however, expresses this aspect intuitively: she describes the need for a (common) uniform subdivision of the intuitive component of the mathematical teaching activity, which is not found in the frame of reference. Intuitively, parts with different subdivisions can be combined. Only the written–symbolic notation of the sum makes a common uniform subdivision necessary. The linking of algorithmic and intuitive components of mathematical teaching activity can be described as contradictory. Thus, the analyzed segment shows an inter-component mathematical inconsistency with regard to the components of mathematical teaching activity.

In the following section, the analyzed segment is considered in the context of the embedded situation. First, the teacher describes a fraction as an evenly divided part, which can be assigned to the intuitive component of mathematical teaching activity according to the reference framework (see green box in Figure 3). Following the segment analyzed above, the teacher comments on making the addends equal—the first sub-step of the addition algorithm. Consequently, the teacher comments on the algorithmic component of mathematical teaching (see yellow box in Figure 3). The (grey) segment analyzed above is thus framed by an intuitive and an algorithmic component of mathematical teaching activity, two components that the segment itself also combines. The analyzed, initially ambiguous interval thus forms the transition from the intuitive to the algorithmic component of mathematical teaching activity. It forms a bridge between the preceding and the subsequent component of mathematical teaching activity.

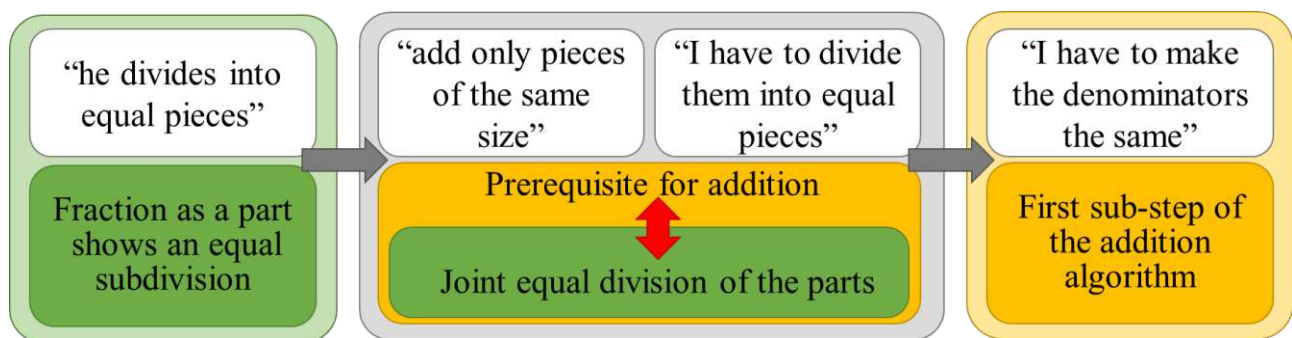


Figure 3: Inter-Component Changes in Segment 00:11:50–00:12:30

FINAL REMARKS

Combination of the described analysis methods enables precise and detailed phases of inter-component changes in the progression of a math unit – insights

A combination of structuring content analysis and adapted context analysis enables a focused analysis of inter-component changes. With the help of the structuring content

analysis, an overview of the course of the teaching unit and the emphasis of the individual components of mathematical teaching activities can be gained. In addition, (potential) changes between the components of mathematical teaching activities can be identified. By adding the context analysis, the framework conditions of the inter-component changes can be considered in detail. In this way, ambiguous intervals can be concretized and located with regard to the uttered components of mathematical teaching activities (see Figure 3).

Phases of inter-component changes in the course of a math lesson

In the lessons observed, changes between the components of mathematical teaching activity occur primarily at the beginning of the lesson and decrease as it progresses. They can be used to identify different phases in the lesson. Some of the inter-component changes at the beginning of the teaching unit can clearly be assigned to one of the components. In this orientation phase, the teacher introduces the mathematical learning content and gives the learners initial access to the various components of mathematical teaching activities. In the subsequent stabilization phase, only potential inter-component changes can be identified. The learning content is narrowed down with regard to the components of mathematical teaching activities used. Finally, in the implementation phase, the teacher only comments on one component of mathematical teaching activity. For the selected teaching content—the addition and subtraction of fractions—this is the algorithmic component. It is assumed that the component of mathematical activity used in the implementation phase is characteristic of the respective lesson content.

Ambiguous intervals can serve as a transition between the components of mathematical teaching activity

The change from the intuitive to the algorithmic component of mathematical teaching activity is interrupted by an ambiguous interval. Through more precise analyses, it can be reconstructed that the ambiguous interval functions as a transition from one component of mathematical teaching activity to the other. It combines the components of mathematical teaching activity of the preceding and subsequent interval. Moreover, the ambiguous coded interval shows a mathematical inconsistency: the expressed contents cannot be located without contradiction of the respective components of the mathematical teaching activity. It is assumed that more mathematical inconsistencies can be found in ambiguous intervals. This must be checked in further analyses.

The identified inter-component changes show potentially challenging situations for teachers. They can serve as a starting point for teacher training to provide teachers with training on these—especially inter-component changes and to sensitize them to inconsistencies. Consequently, teachers are supported in presenting inter-component changes in a mathematically consistent way and thus stimulating meaningful learning processes in learners.

References

- Copur-Gencturk, Y. (2021). Teachers' conceptual understanding of fraction operations: Results from a national sample of elementary school teachers. *Educational Studies in Mathematics*, 107(3), 525–545. <https://doi.org/10.1007/s10649-021-10033-4>
- Fischbein, E. (1994). The Interaction between the formal, the algorithmic, and the intuitive components in a mathematical activity. In R. Biehler, R. W. Scholz, R. Sträßer, & B. Winkelmann (Eds.), *Mathematics education library: Vol. 13. Didactics of mathematics as a scientific discipline* (pp. 231–245). Kluwer Academic Publishers.
- Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., & Hachfeld, A. (2013). Professional competence of teachers: Effects on instructional quality and student development. *Journal of Educational Psychology*, 105(3), 805–820. <https://doi.org/10.1037/a0032583>
- Mayring, P. (2014). *Qualitative content analysis: theoretical foundation, basic procedures and software solution*. Social Science Open Access Repository. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-395173>
- Möller, V. (in press-a). Mathematische Konsistenz – individuelle mathematische Lehrtätigkeit zwischen Intuition, Algorithmus und Formalem [Mathematical consistency - individual mathematical teaching activities between intuition, algorithm and formality]. In Conference Organizers (Eds.), *Beiträge zum Mathematikunterricht 2025*. Münster: WTM.
- Möller, V. (in press-b). Multimodale Rekonstruktion individueller Lehrkonzepte von Lehrpersonen: instruktionale Erklärsituationen im authentischen Mathematikunterricht [Multimodal reconstruction of individual teaching concepts of teachers: Instructional explanatory situations in authentic mathematics lessons]. Waxmann.
- Prediger, S. (2008). Discontinuities for mental models – A source for difficulties with the multiplication of fractions. In D. de Bock, B. D. Sondergaard, & B. A. Gómez (Eds.), *Proceedings of ICME-11: TSG10 Research and development in the teaching and learning of number systems and arithmetic: 11th International Congress on Mathematical Education, Monterrey, Mexico* (pp. 29–37). Katholieke Universiteit Leuven.
- Vogel, R. (2017). „wenn man da von oben guckt sieht das aus als ob ...“ – die ‚Dimensionslücke‘ zwischen zweidimensionaler Darstellung dreidimensionaler Objekte im multimodalen Austausch [The 'dimensional gap' between two-dimensional representation of three-dimensional objects in multimodal communication]. In M. Beck & R. Vogel (Eds.), *Geometrische Aktivitäten und Gespräche von Kindern im Blick qualitativen Forschens: Mehrperspektivische Ergebnisse aus den Projekten erStMaL und MaKreKi* (pp. 61–75). Waxmann.
- Weinert, F. E. (2014). Vergleichende Leistungsmessung in Schulen -eine umstrittene Selbstverständlichkeit [Comparative performance measurement in schools - a controversial matter of course]. In F. E. Weinert (Ed.), *Pädagogik. Leistungsmessungen in Schulen* (pp. 17–32). Beltz.
- Wittwer, J., & Renkl, A. (2008). Why instructional explanations often do not work: A framework for understanding the effectiveness of instructional explanations. *Educational Psychologist*, 43(1), 49–64. <https://doi.org/10.1080/00461520701756420>

DIDACTIC AND DISCIPLINARY KNOWLEDGE OF ELEMENTARY SCHOOL TEACHERS IN COSTA RICA: A STUDY OF STATISTICS

José Luis Morales Reyes¹, and Ana Lucía Alfaro Arce²

Pontificia Universidad Católica de Valparaíso¹, Universidad Nacional de Costa Rica²

This paper examines the didactic-mathematical knowledge expertise demonstrated by a collective of elementary school teachers when confronted with solving tasks related to statistics. The study is based on the didactic suitability aspects of the Onto-Semiotic Approach and employs an exploratory and mixed approach, through the analysis of the answers given by 17 teachers. The results indicate that the teachers encounter difficulties in activating and recognizing the necessary knowledge, skills, and attributes essential for accurate problem resolution. Furthermore, the research underscores the necessity for a dedicated qualification framework to direct the foundational mathematical knowledge requisite for the training of elementary teachers with a generalist background.

INTRODUCTION

In Costa Rica, elementary education encompasses students aged 7 to 12. Following a curricular reform implemented in 2012, the teaching of statistics and probability is now required at all educational levels, from elementary through high school. In accordance with the guidelines set by the Ministerio de Educación Pública (MEP), this integration aims to foster a culture of managing and interpreting both quantitative and qualitative information from students' environments. Additionally, it seeks to deepen their understanding of random phenomena within their context (MEP, 2012).

On the other hand, considering that in Costa Rica there are still no established guidelines regarding the profile, number of years and knowledge that teacher training plans should contemplate then there is a great diversity of them among universities (Alpízar & Alfaro, 2019). For this reason, a qualifications framework was established in 2021 that aims to homogenize the knowledge that elementary school teachers should receive in their professional studies. However, it is not yet mandatory, and it is generalist, as it does not specifically detail the learning outcomes in any of the areas of knowledge (Mathematics, Spanish, Social Studies, Science) of basic general education (Alfaro & Morales, 2023).

Therefore, the objective of this research is to investigate the depth of knowledge that elementary teachers have about the mathematical domains covered by the official curriculum of the MEP (20XX), particularly in the topics of Statistics; since, as Godino et al. (2017) points out, teachers have to know the school mathematics of the

educational level where they teach, but must also be able to articulate that knowledge with that corresponding to some subsequent levels.

Based on the described problem, this type of study is a key element both for expanding research in this field and for assisting in decision-making in the (re)formulation of initial training plans.

THEORETICAL APPROACH

There are several models that attempt to provide an approach to the type of knowledge that a teacher should have to teach mathematics, regardless of the level of instruction (Scheiner et al., 2019). In this research, we used the model of mathematics teachers' knowledge known as Didactic-Mathematical Knowledge (DMK), which is based upon theoretical assumptions and theoretical–methodological tools of Onto-Semiotic Approach to mathematical cognition and instruction (Godino et al., 2007), this theoretical system aims to “address the epistemological, ontological, semiotic, cognitive, and instructional dimensions involved in mathematics education in a unified manner” (Barquero et al., 2022, p. 241).

Specifically, we worked with didactic suitability (Pino-Fan et al., 2015). The didactic suitability of a formal instructional process is defined as the degree to which such process or part of it meets certain criteria, previously defined, that allow qualifying it as adequate to achieve the adaptation between students' learning and the intended institutional meanings (what is to be taught) considering the environment where such process takes place (Godino et al., 2016). Within didactic suitability according to Godino et al. (2007), six criteria are established: epistemic, cognitive, interactional, mediational, emotional, and ecological (see Figure 1).

These criteria are defined by Godino et al. (2006) and Font et al. (2010) as followed:

- Epistemic suitability is the didactic-mathematical knowledge about the content itself, that is, the way in which the mathematics teacher understands and knows mathematics. Considering as a reference the prescribed curriculum; that is, the institutional mathematics that have been transposed into the curriculum and in general the mathematical content studied.
- Ecological suitability is the degree of adaptation of the study process to the educational project of the center, the curricular guidelines, the social environment, relations with other subjects, etc.
- Mediational suitability is the degree of availability and adequacy of the material and temporal resources necessary for the development of the teaching-learning process.
- Interactional appropriateness is the degree to which the interactions allow the identification and resolution of conflicts of meaning and doubts to favor autonomy in learning.
- Cognitive suitability, assesses, before starting the instruction process, if what is to be taught is at a reasonable distance from what the students know, and

after the process, if the learning acquired is close to what was intended to be taught.

- Affective appropriateness is the degree of involvement (interest, motivation) of the students in the instructional process.

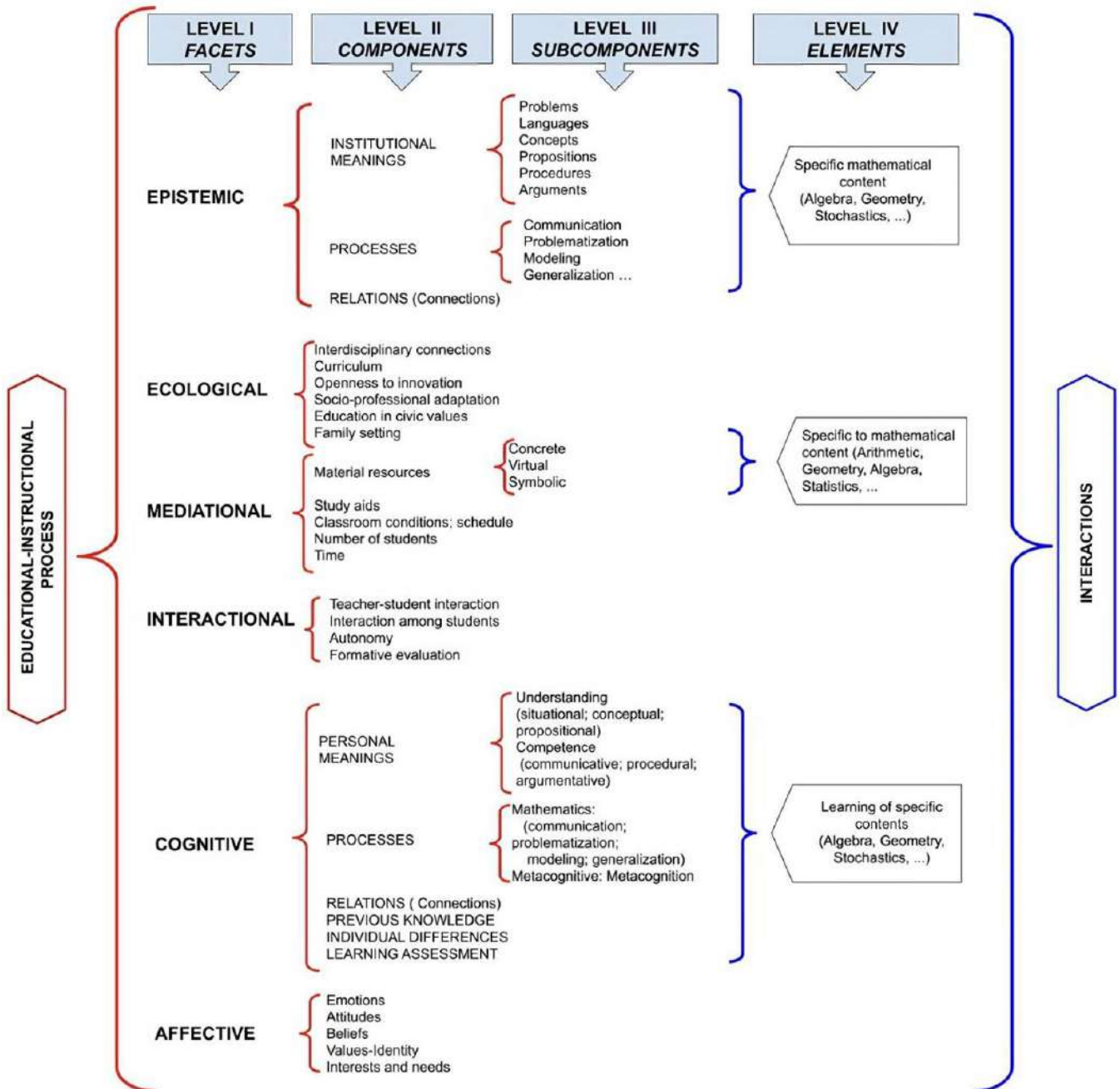


Figure 1: Facet and components of didactic suitability (Godino et al., 2023)

METHODOLOGY

The research is exploratory, with a mixed-methods approach. The analysis of the variables is both quantitative, focusing on correct or incorrect solutions, and qualitative, considering various types of arguments, justifications, errors, difficulties,

and other factors involved in solving the problem situations presented in the questionnaire items. The entire study was conducted in Spanish.

Within this research, we worked with a sample, chosen at convenience, of teachers in service, in different public and private elementary schools in two states of Costa Rica. Data collection was obtained throughout 2020, the participation of 17 elementary school teachers was achieved, representing five different institutions (two private and three public). Eight of the teachers had received courses in statistics and probability as part of their university education, and of these, three indicated that this knowledge was sufficient for their work.

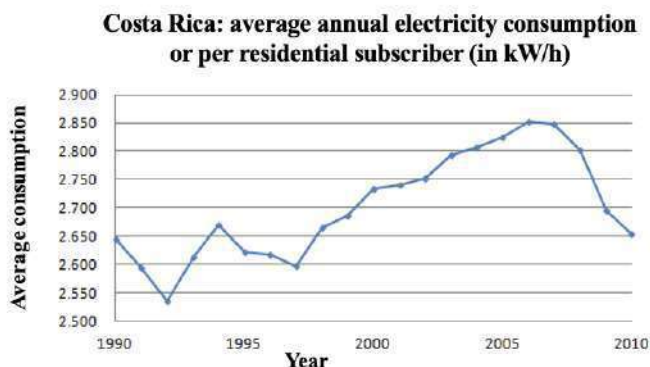
In addition, a review of the official elementary school curriculum was previously carried out to determine the minimum mathematical knowledge required by these teachers, i.e., the knowledge that their students should have at the end of elementary school. Thus, the questionnaire was designed based on this program, and items related to other didactic suitability were added. It should be noted that the questionnaire focused on the competencies of the second cycle (students from 10 to 12 years of age), which corresponds to the last years of primary education and where the transition from intuitive and introductory to more formal elements begin.

The instrument for obtaining the data was based on the model of categories of DMK and on the methodology it proposes; first, a mathematical task is chosen that leads the teacher to consider different aspects of statistics through the solution of the task or problem situation posed; and then evaluation questions or proposals for activities are formulated that include the different facets of the teacher's knowledge that are to be analyzed.

Thus, by means of this open-response instrument, it is possible, from the productions expressed in the solution of the different questions that make up the questionnaire, to obtain some empirical indicators of the didactic and statistical knowledge of the participants, allowing, in a certain way, the evaluation of such knowledge, an aspect that is not always possible to achieve by simple observation or survey with closed questions.

The questionnaire was validated in several stages (with international and national teachers with training in Mathematics Education; and through pilot testing with practicing teachers in elementary education). The complete research instrument is composed of thirteen general questions and six disciplinary and didactic questions (five on statistics and one on probability). Due to space limitations and as an example, in this article we show only one of the statistical items in Figure 2 (see Alfaro and Morales (2023) for another more detailed statistical item with student responses).

Ms. Francisca extracted the attached graph from a national newspaper and asked her fifth grade students to analyze the information shown regarding average annual electricity consumption per residential subscriber in kilowatt hours (kW/h).



Some of Ms. Francisca's students gave the following interpretations of the information they read from the graph.

Dariana: the average electricity consumption decreased from 2006.

Franco: the maximum electricity consumption presented in a home was 2850 kW/h.

Ariana: total electricity consumption among residential subscribers was approximately the same in 1990 and 2010.

Figure 2: Example of one of the situations included in the questionnaire.

For this situation, four questions were asked, related to the facets of didactic suitability. Table 1 shows the questions and their respective classification.

Questions items	Didactic suitability facet
a) Make interpretations of the information you read from the graph.	Epistemic facet
b) Which one(s) of Ms. Francisca's students have given a correct interpretation of the information shown in the graph? Justify your answer.	Epistemic and Cognitive facets (knowledge of how students' reason)
c) What processes (reasoning and argumentation, problem solving, communicating, connecting, representing) could be enhanced in fifth grade students by discussing the information in the graph? Briefly describe how you would conduct an activity to enhance the most processes.	Mediational, interactional, epistemic, and affective (motivation) facet
d) Briefly describe how the information in the graph could be used to discuss with students the usefulness of statistics in everyday life?	Ecological facet

Table 1: Items and its relation to didactic suitability facets

RESULTS

The distribution of the results according to the degree of success for each of the questions is presented in Table 2.

Item	Incorrect or no answer	Partially correct	Correct	Total
a	15	1	1	17
b	10	7	0	17
c	13	3	1	17
d	6	7	4	17

Table 2: Number of people by degree of accuracy of content knowledge

The epistemic phase refers to the teacher's understanding of the relevant statistical content, which requires consideration of both the diversity of meanings and the variety of objects and processes involved. In the activity discussed in this paper, the key concept of average consumption (item a) was not identified in 15 out of 17 cases. Moreover, participants struggled with interpreting the data: many focused only on identifying the minimum or maximum values of the graph, or highlighted isolated data points, without recognizing trends. Some participants even failed to grasp the concept of interpretation altogether, as their responses did not address the question. This suggests that, for many, the ability to perform a basic level of interpretation—such as reading information from a graph—was not achieved, as illustrated in Figure 3.

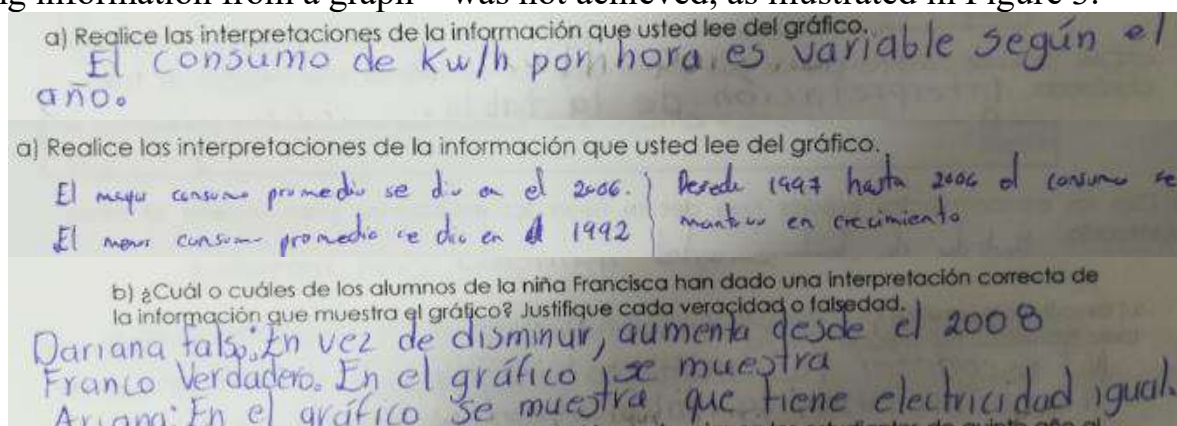


Figure 3: Examples of the type of responses obtained in the epistemic facet.

Note. English translation: (1a) Kw/h per hour consumption varies from year to year. (2a) The highest average consumption was in 2007. The lowest average consumption was in 1992. From 1997 to 2006 consumption showed growth. (3b) Dariana: false. Instead of decreasing, it has increased since 2008. Franco: true. The graph shows it. Ariana: the graph shows that they have equal electricity consumption.

The results indicate that, in general, teachers are unaware of the mathematical processes—referred to as competencies in other curricula—that could be promoted through the activity, such as reasoning and argumentation when justifying the correctness of an answer, as well as communication skills for expressing interpretations

derived from graphic analysis. This lack of awareness persists despite the fact that these processes are fundamental in the theoretical foundations of the Costa Rican curriculum.

Additionally, teachers were generally able to provide satisfactory responses regarding how to use the given context to discuss the relevance of statistics in daily life, an aspect that is central to elementary mathematics education and aims to foster positive attitudes toward the subject.

CONCLUSIONS

This study explored initial and partial aspects of the mathematical didactic knowledge of elementary school teachers in Costa Rica. Although it was exploratory and descriptive in nature, a thorough analysis of teachers' knowledge was conducted across each of the facets of didactic suitability: content (epistemic), students (affective and cognitive), teaching (interactional), and curriculum (ecological).

The results suggest that their specialized knowledge for teaching statistics is insufficient across all dimensions. No justifications or arguments were provided, even though the instrument emphasized this point. The results reveal that participating teachers struggled to identify and mobilize the necessary content and knowledge to provide correct solutions to the posed situations (epistemic facet). There was also a lack of understanding of the evaluated concepts and their meanings, as teachers could not describe students' cognitive processes, learning difficulties, or errors, nor suggest strategies to support them (affective and cognitive facets). Additionally, there was limited knowledge of the interactional and mediational facets; while some correct answers were provided, the justifications were overly general.

In conclusion, there is an urgent need for universities to develop training programs that enhance the mathematical knowledge of elementary school teachers, thereby contributing to the improvement of statistical education in Costa Rica. Additionally, given the current diversity of university programs in elementary education (Alpízar & Alfaro, 2019), this study highlights the necessity of establishing a specific qualification framework. Such a framework would guide teacher educators in identifying the minimum mathematical knowledge that must be covered in university curricula, in alignment with the official curriculum that future generalist teachers will be expected to teach.

Acknowledgements

The first author was supported by the Chilean Government Doctoral Scholarship ANID [folio: 21210589].

References

- Alfaro, A., & Morales, J. L. (2023). Conocimientos didácticos y estadísticos de un grupo de docentes de II ciclo de Educación Primaria en Costa Rica (Didactic and statistical knowledge of a group of elementary school teachers in Costa Rica). *Repertorio Científico*, 26(2), 76–89. <https://doi.org/10.22458/rc.v26i2.5076>

- Alpízar, M. & Alfaro, A. (2019). College education of elementary school teachers: the case of mathematics. *Uniciencia*, 33(2), 110-154. <https://doi.org/10.15359/ru.33-2.8>
- Barquero, B., Batanero, C., Blanco, T., Bosch, M., Camacho-Machín, M., Cañadas, M., Castro, E., Chamoso, J. M., Contrera, L., Gea, M., Godino, J.D., Martín-Molina, V., Moreno, M., & Wilhelmi, M. R. (2022). Spanish research on Mathematics Education. In C. Fernández, S. Llinares, A. Gutiérrez, & N. Planas (Eds.), *Proc. 45th Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 1, pp. 223-254). PME.
- Godino, J.D., Batanero, C. & Font, V. (2007). The onto-semiotic approach to research in mathematics education. *ZDM Mathematics Education* 39, 127–135. <https://doi.org/10.1007/s11858-006-0004-1>
- Godino, J.D., Batanero, C., Font, V., Contreras, A. & Wilhelmi, M. (2016). *The theory of didactical suitability: networking a system of didactics principles for mathematics education from different theoretical perspectives*. 13th International Congress on Mathematical Education.
- Godino, J. D., Batanero, C., & Burgos, M. (2023). Theory of didactical suitability: An enlarged view of the quality of mathematics instruction. *Eurasia Journal of Mathematics, Science and Technology Education*, 19(6), 1-20. <https://doi.org/10.29333/ejmste/13187>
- Ministerio de Educación Pública (MEP). (2012). *Programas de Estudio en Matemáticas para la Educación General Básica y el Ciclo Diversificado [Programs of Study in Mathematics for General Basic Education and the Diversified Cycle]*. San José, Costa Rica: autor.
- Pino-Fan, L., Assis, A., & Castro, W. F. (2015). Towards a methodology for the characterization of teachers' didactic-mathematical knowledge. *Eurasia Journal of Mathematics, Science & Technology Education*, 11(6), 1429–1456. <https://doi.org/10.12973/eurasia.2015.1403a>
- Scheiner, T., Montes, M., Godino, J.D., Carrillo, J., & Pino-Fan, L. (2019). What Makes Mathematics Teacher Knowledge Specialized? Offering Alternative Views. *International Journal of Science and Mathematics Education*, 17, 153–172. <https://doi.org/10.1007/s10763-017-9859-6>

IMPLICATIONS FOR TEACHER TRAINING OF INCONSISTENCIES BETWEEN UNIVERSITY AND SECONDARY TEXTBOOKS REGARDING THE MULTIPLICATION PRINCIPLE

Jonathan Morcos and Alejandro S. González-Martín

Département de Didactique, Université de Montréal, Canada

This paper examines and compares the rationales provided for the multiplication principle (MP) in combinatorics, as presented in secondary school textbooks and university-level textbooks used for training prospective secondary mathematics teachers (PSMTs). Numerous inconsistencies are identified, including differences in the scope of the MP's application. We argue that the university textbooks provide weak rationales that are misaligned with secondary teaching practices. Interviews with PSMTs support our findings, revealing how these inconsistencies and the absence of clear rationales may affect the teaching of the MP in secondary schools.

INTRODUCTION AND RESEARCH PROBLEM

Combinatorics, a branch of mathematics concerned with discrete sets of objects, can be a difficult subject for students of any age (Lockwood et al., 2020). In this paper, we focus on one challenging concept in particular: the multiplication principle (MP). As its name implies, the MP is used to solve combinatorics tasks that involve multiplication, particularly in counting processes of two or more independent stages. Lockwood et al. (2017) analysed 64 university textbooks' statement of the MP. They identified a broad variety of wordings: "the MP is much more nuanced than instructors and students perhaps give it credit for. Given its foundational place in combinatorial enumeration, we need to help students focus more on understanding the details of the MP" (p. 411). Lockwood and Purdy's (2019) study with undergraduate students reveals that students have difficulty precisely defining and describing actions associated with the operation of multiplication in combinatorics tasks.

We believe this is highly relevant to teacher training, as teachers must be able to articulate clear explanations and justifications with respect to the MP. However, teacher training in combinatorics remains underexplored (Lockwood et al., 2020), and, to our knowledge, there is scant research examining the use of the MP in the training of prospective secondary mathematics teachers (PSMTs). In one rare study, Lamanna et al. (2023) identify the MP (referred to as the product rule) as a prominent strategy used by PSMTs and secondary students in solving combinatorics tasks, emphasising the need to study this technique and its accessibility to PSMTs.

Tertiary mathematics courses are a standard component of teacher training programs. PSMTs are expected to master the mathematical content they will teach and draw on their tertiary education to inform their teaching practices (Biza et al., 2022). However,

numerous inconsistencies between secondary and tertiary mathematics instruction—including differences in content, mathematical practices, teaching methods, and other factors—hinder teachers' ability to apply the combinatorics knowledge gained at university in their classrooms. As a result, PSMTs often view these courses as unproductive and irrelevant (Biza et al., 2022). We aim to explore this potential discontinuity (Klein, 1908/2016), particularly regarding the MP.

In this paper, we seek to better understand how university programs prepare PSMTs to teach the MP. Drawing on data from the first author's ongoing master's research, we address the following research questions: *What rationales related to the MP are presented in tertiary mathematics courses for PSMTs? How do these rationales compare to the use of the MP at the secondary level? If inconsistencies can be identified, what are the implications for PSMTs?*

THEORETICAL FRAMEWORK

This paper examines the MP in a PSMT university program and in secondary education. Rather than focusing on the individual competencies of PSMTs, we analyse how the MP is addressed at an institutional level, both in PSMT training and in secondary mathematics instruction. Our theoretical framework is based on the Anthropological Theory of the Didactic (ATD), which views tertiary mathematics and secondary teaching as distinct didactic *institutions*, each with its own characteristics (Chevallard & Bosch, 2020). According to ATD, scholarly content is adapted to meet the needs of different institutions through the process of didactic transposition, which can result in discrepancies in how the “same” content is presented in different institutions (Castela, 2017). We seek to explore how the MP is organised in relation to other combinatorics content within tertiary mathematics and secondary teaching, identifying consistencies and inconsistencies across these two institutions.

ATD characterises institutionally situated content through praxeologies (Chevallard & Bosch, 2020), which consist of two blocks. The first block, *praxis*, comprises types of tasks and the techniques used to solve them. The second, *logos*, encompasses the theoretical elements that justify the praxis. This includes rationales, referred to as *technologies*, and elements of theory that ensure mathematical rigour. Our aim is to analyse and compare the logos associated with the MP in university mathematics and secondary teaching, as it forms the basis for a rigorous rationale. We also consider certain elements of the praxis, given its intrinsic connection to the logos (e.g., the justifications of techniques). We therefore refine our research questions as follows: *What logos and associated praxis related to the MP are present in the university training of PSMTs? How does this logos relate to praxeologies in secondary teaching? If inconsistencies exist, what are the potential implications for PSMTs?*

METHODOLOGY

Our research, conducted at Université de Montréal in Québec, Canada, is based on two phases of data collection. In the first phase, we analysed two university textbooks

(UTs) used in PSMT courses that cover combinatorics: *Discrete mathematics* and *Probability*. We also analysed three secondary mathematics textbooks (STs) approved by the Québec Ministry of Education, along with their teacher's guides. Full references for these textbooks are provided [here](#). We applied open coding methods (Saldaña, 2013) to investigate MP-related praxeologies in the sections of each textbook's chapter covering combinatorics. Tasks involving the MP (explicitly or implicitly) were identified, and the accompanying rationales were analysed, revealing consistencies and inconsistencies in how the MP is addressed in tertiary and secondary mathematics (Morcos & González-Martín, in press). This led to conjectures about potential difficulties arising from using these textbooks. This paper aims to further explore these observations by examining how these inconsistencies might hinder PSMTs' understanding and teaching of the MP.

To triangulate these findings, we conducted interviews with PSMTs in a second phase of data collection. This paper incorporates insights from the first two interviews only, as this phase remains ongoing. The two participants, P1 and P2, are in their last year of university, with only their final practicum left to complete. During the interviews, the participants solved combinatorics tasks selected from the analysed UTs and STs. They were asked to justify their solving techniques and connect their approaches and rationales to either their university training or secondary teaching experiences. The interviews were transcribed for analysis, with excerpts translated into English.

For both phases, the first author conducted an initial data analysis, which was verified by the second author. Minor disagreements were resolved as needed.

RESULTS AND DISCUSSION

The MP in university textbooks

In both of the UTs, the MP is presented as a fundamental element of combinatorics. In UT1, it is introduced as the first theorem of the chapter, referred to as the *fundamental principle* and deemed “essential” (p. 2) for subsequent content. In UT2, the MP appears in the first subsection of the combinatorics chapter: *Fundamental principles of counting*. This rationale emphasises the crucial role of multiplication in solving combinatorics tasks, with 16 of 47 tasks (34.0%) in UT1 and 21 of 40 tasks (52.5%) in UT2 relying on the MP for their solutions, either implicitly or explicitly. Both UTs define the MP similarly, consistent with Lockwood et al.'s (2017) inventory:

UT1: “Suppose two experiments need to be conducted. If experiment 1 can produce any one of m outcomes, and for each of these there are n possible outcomes for experiment 2, then there are $m \cdot n$ outcomes for the two experiments combined.” (p. 2)

UT2: “Assume that a procedure can be divided into two tasks. If there are n_1 ways to complete the first task and n_2 ways to accomplish the second task once the first is completed, then there are $n_1 \cdot n_2$ ways to carry out the procedure.” (p. 220)

Once it is introduced, both textbooks provide weak mathematical justifications for the MP. In UT2, the text simply states: “The multiplication principle applies when a

procedure consists of different consecutive tasks. [...] Examples 4 and 5 show how to use the multiplication principle” (p. 220). In UT1, the justification is based on counting outcomes organised in a two-dimensional $m \times n$ matrix, explained as:

In this table, each result is noted (i, j) if the first experiment produces the i^{th} outcome and the second experiment the j^{th} . The total possible results consist of m rows of n elements each, demonstrating the stated result. (p. 2)

This rationale inadequately justifies the need for multiplication in contexts requiring the MP. For example, while it mentions task independence, UT1 fails to elaborate on potentially confusing nuances, such as “[...] the sets of options at each stage need not be independent, but the cardinality of the sets must be independent” (Lockwood & Purdy, 2019, p. 232). Additionally, relying on a matrix to justify the MP is not easily transferable to the generalised version for r experiments: no straightforward visual aid, such as a two-dimensional matrix, can be provided for three or more multiplication steps. Neither UT1 nor UT2 offers a justification for the *generalised fundamental principle* (see Figure 1). This reveals a lack of rationale in both UTs for two-stage and multi-stage multiplication procedures in combinatorics.

1.2.2 The generalised fundamental principle

When there are more than two experiments to carry out, the fundamental principle can be generalised as such:

Theorem 1.2

If there are r experiments to carry out, such that the first can produce any one of n_1 results, if for each of them there are n_2 possible results for the second experiment, if for each result from the first two experiments there are n_3 possible results for the third experiment, and so on, then there are a total of $n_1 \cdot n_2 \cdot \dots \cdot n_r$ results for the r experiments taken together.

1.2.3 Examples of the use of the fundamental principle

Figure 1: Presentation of the MP in r -stages experiments in UT1 (pp. 2–3)

Both textbooks prioritise applying the theorem over constructing it through detailed mathematical rationales. This reliance on formulas aligns with other findings about the textbooks’ rationale for MP-related tasks: the MP is used as the technology supporting multiplication. In all 21 tasks in UT2 involving multiplication, the MP serves as the explicit rationale (“By the multiplication principle...”) despite the fact that no initial rationale is given for this technique. In UT1, the MP gradually becomes implicit, a shift often associated with the routinisation of techniques (Chevallard & Bosch, 2020), potentially reinforced by the reliance on formulas in the combinatorics chapters of both UTs. Moreover, this approach suggests that the MP may be applied automatically to any combinatorics task requiring multiplication, without questioning the validity of the technique or fostering a deeper understanding of the underlying principles.

Thus, despite the MP’s central role in combinatorics solving techniques in both textbooks, the justifications provided fail to promote a deep understanding of important mathematical concepts. As a result, undergraduate students relying on these textbooks may struggle to justify the use of the MP in combinatorics tasks. This could present a significant barrier for PSMTs, as the absence of robust rationales may lead to

difficulties teaching the MP. Our interviews shed light on this issue. In the next section, we compare the UTs with the STs and discuss the interview data.

The MP in secondary textbooks and interviews with PSMTs

The STs introduce combinatorics in their chapters on probability. As in the UTs, the MP is frequently applied: it appears in 16 of 28 tasks (57.1%) in ST1, 14 of 23 tasks (60.9%) in ST2, and 12 of 39 tasks (30.8%) in ST3. However, the presentation of the MP varies significantly between these textbooks. In ST1, the only explicit rationale for multiplication pertains to probabilities: “The probability of an elementary event in a multi-stage random experiment equals the product of the probability of each intermediate event at each stage” (p. 342). In ST2, the MP is briefly mentioned at the end of the chapter, described as a technique appropriate for “certain situations involving regularities” (p. 665), but no formal definition is given. Instead, the rationale gives an example that applies the MP to a task without replacement and with ordered outcomes. The MP is, however, used earlier in ST2 to solve tasks both with and without these constraints. Thus, the praxis and logos for the MP in ST2 remain vaguely defined.

ST3 places the greatest emphasis on the MP as a technique. Its teacher’s guide states: “Counting methods are based on the multiplication principle. The aim of this section is to remind pupils of this important principle [...]” (p. 354). Two definitions are provided in the combinatorics chapter: 1) “[...] if there are m ways to do one task and n ways to do another, then there are $m \cdot n$ ways to do both tasks in a given order” (p. 357); 2) “The total number of possibilities for a multi-step random experiment equals the product of possibilities for each step” (p. 371). The first statement closely resembles UT2’s, but it appears only in the teacher’s guide, in the solution for one specific task. In contrast, the second definition, given later as an explicit explanation in the textbook, is more prominent, likely encouraging its use over the first one.

In all three STs, the MP is associated with tree diagrams. This other combinatorics technique involves drawing branches representing possible outcomes at each step, with the total number of branches indicating the total number of outcomes. The teacher’s guides for all three textbooks make connections between the two techniques:

ST1: “Some pupils may find it difficult to calculate the number of codes possible. If necessary, ask them to construct a probability tree.” (p. 333)

ST2: “Teachers can help pupils [struggling to determine possibilities (e.g., 52×51)] by inviting them to start building a tree diagram to visualise all possibilities and understand the multiplication principle in combinatorics.” (p. 537) Later: “[...] pupils may use [...] visual representations (tree diagrams, grids, etc.), but [...] will soon realise these have limits [and] must reason using the multiplication principle.” (p. 539)

ST3: “[Pupils] will soon see they cannot build [the complete tree diagram], because there are too many results. However, even a partial tree diagram can help them ‘see’ the required operation to determine the number of [possible results].” (p. 355)

Three observations are noteworthy. First, the link between both techniques is emphasised in all three STs but is absent from the UTs. P1 highlighted this gap:

P1: I know there's a calculation that comes from all of this, from the tree diagram, but I can't recall the formula. So, I'll just do it with the tree diagram [...] It's the easiest way to see it. It's a multiplication between each branch. [...] That's how we did it in secondary school.

Second, while UT2 lacks visual aids, UT1 and the STs note their usefulness in justifying multiplication. However, there is an inconsistency in their representations: UT1 employs a matrix, while the STs use tree diagrams. This misalignment may render the tools provided to PSMTs unsuitable for secondary teaching:

P1: [...] I don't think [the probability course] will be useful, at all [...] I never saw a single tree diagram during my time at university.

P2: [At university, I didn't learn] what to use when. We made extensive use of formulas, [...] but for teaching [...], diagrams or other graphical representations that help find the answers and that are not purely analytical, I think that is missing [from the university courses].

Third, the UTs and STs differ epistemologically regarding the MP: the UTs build *from* the MP, whereas the STs build *to* it. The UTs introduce the MP early as underpinning other combinatorics theorems and formulas, such as permutations and combinations, a pattern noted by other researchers (Lockwood et al., 2015, in Lockwood & Purdy, 2019). In contrast, the STs present and use the MP in tasks where alternatives, such as the tree diagram, are impractical. This approach diverges from two ideas presented in the UTs: 1) the MP is a fundamental principle of counting; 2) the MP applies broadly to tasks involving multiplication. For PSMTs, these inconsistencies may hinder the transfer of knowledge from university to secondary teaching. Our analysis also indicates that the weak rationales in the UTs risk undermining PSMTs' deep understanding of the MP, while the weak rationales in the STs may fail to compensate for this gap, leaving teachers with limited tools to address their pupils' difficulties. This challenge was evident in the interviews: Figure 2 illustrates P1's struggles to justify their reasoning, despite successfully solving the given task with no apparent difficulty.

Task: "How many different organisations of the letters in the word PEPPER are there"	
Int:	Here, the multiplication, why did you multiply?
P1:	Okay, why multiply? Because... You have the possibilities for the Ps, at first, that's 6 choose 3. Then I have to, I have to, I can't add because if I add it just doesn't make sense because I just, like, add the, heum... Wait, I just want to clear my head. So it's 6 choose 3, I place my Ps, then I place my Es. So, once they're placed, I can add an additional constraint, so its like a multiplication. I know that I'm saying 'add', but like in my head it's pretty clear, like, I add the Es so it's like a multiplication because I add an additional constraint so I have to multiply the two. That's what I would say.

Figure 2: Excerpt from P1's interview focusing on the rationale for multiplication

Lastly, the placement of combinatorics within the STs' chapters on probability affects the MP's praxis. The MP is applied in tasks both with and without order and repetition, yet no clear rationale distinguishes these cases. In ST1, the MP is explicitly addressed only in the section *Multi-stage random experiments with or without repetition*, where the focus is on probabilities rather than on counting outcomes. In ST2, rationales for

the MP and the product of probabilities are presented at the end of the chapter, after the resolution of all tasks, while ST3 links the MP to probability calculations for compound events involving independent steps. However, the probability of an event is calculated by dividing favourable and total outcomes. This approach allows for using the MP to consider ordered outcomes in both steps of the division, regardless of initial constraints in the context of the tasks, as this order ultimately “cancels out”. This conflates the use of the MP in ordered and unordered contexts. It also diverges from the UTs, where tasks with unordered outcomes usually involve using the formula for combinations. These inconsistencies emerged in the interviews. P2 struggled in two tasks to discern whether the MP counts ordered or unordered outcomes, reflecting difficulties identified among undergraduates by Lockwood & Purdy (2020). Moreover, both P1 and P2 applied the MP to a task with unordered outcomes taken from an ST, while noting that at university, the formula for combinations is preferred. This suggests that at the university level, the MP is associated with tasks where order matters, whereas at secondary school it is applied for both ordered and unordered outcomes.

Our data suggests that the differences in the way the UTs and STs present their content influence the understanding of the MP and may also impact the future teaching practices of PSMTs. Further analysis of the institutional division of content—order versus unordered, with repetition versus without repetition, combinatorics versus probability—could provide additional insights.

CONCLUSION

This paper compares praxeologies related to the MP in the university training of PSMTs and in secondary school settings. Our findings reveal a lack of robust rationales in UTs, which may impede PSMTs’ ability to provide appropriate justifications and explanations to their pupils. Furthermore, visual aids and task types involving the MP in UTs do not align with the content in STs, while the latter offer little to compensate for absent rationales.

Our theoretical framework facilitated the analysis of practices associated with the MP through the lenses of praxis and logos, examining the presence or absence of rationales (technologies) justifying the MP’s application and the tasks and solving techniques that invoke the MP. This approach uncovered institutional inconsistencies that PSMTs may face when transferring knowledge gained at university to secondary classrooms. Our interview data corroborate our textbook analysis findings, suggesting that classroom practices may mirror the textbooks’ approaches. Our results complement Lockwood et al.’s (2017) analysis of the MP in textbooks by considering more than the MP statement, revealing unclear rationales as a potential root of university and secondary students’ difficulties with the MP. Our study also suggests that differing practices at university and secondary levels may exacerbate these difficulties. Considering that the study of combinatorics for teachers is an underexplored research area (Lockwood et al., 2020), we believe our findings contribute significantly to this field. The next steps of this research will serve as a source for future publications.

ACKNOWLEDGMENTS

Funded by the Fonds de recherche du Québec, Secteur Société et culture (#343978) and the Social Sciences and Humanities Research Council (#766-2024-1117).

References

- Biza, I., González-Martín, A. S., & Pinto, A. (2022). “Scaffolding” or “filtering”: a review of studies on the diverse role of calculus courses for students, professionals and teachers. *International Journal of Research in Undergraduate Mathematics Education*, 8(2), 389–418. <https://doi.org/10.1007/s40753-022-00180-1>
- Castela, C. (2017). When praxeologies move from an institution to another: an epistemological approach to boundary crossing. In R. Göller, R. Biehler, R. Hochmuth, & H.-G. Rück (dir.), *Proceedings of the KHDM conference: Didactics of mathematics in higher education as a scientific discipline* (p. 418–425). Universitätsbibliothek Kassel.
- Chevallard, Y., & Bosch, M. (2020). Anthropological Theory of the Didactic (ATD). In S. Lerman (ed.), *Encyclopedia of Mathematics Education* (pp. 53–61). Springer. https://doi.org/10.1007/978-3-030-15789-0_100034
- Klein, F. (1908/2016). Elementary mathematics from a higher standpoint: Volume I: Arithmetic, Algebra, Analysis. Springer. <https://doi.org/10.1007/978-3-662-49442-4>
- Lamanna, L., Branchetti, L., & Bolondi, G. (2023). Combinatorial reasoning in problem solving strategies: a pilot study with pre-service teachers. In P. Drijvers, C. Csapodi, H. Palmer, K. Gosztonyi, & E. Konya (eds.), *Proceedings of the Thirteenth Congress of the European Society for Research in Mathematics Education (CERME13)* (pp. 964–971). Alfréd Rényi Institute of Mathematics et ERME.
- Lockwood, E., & Purdy, B. (2019). Two undergraduate students’ reinvention of the multiplication principle. *Journal for Research in Mathematics Education*, 50(3), 225–267. <https://doi.org/10.5951/jresmetheduc.50.3.0225>
- Lockwood, E., & Purdy, B. (2020). An unexpected outcome: students’ focus on order in the multiplication principle. *International Journal of Research in Undergraduate Mathematics Education*, 6(2), 213–244. <https://doi.org/10.1007/s40753-019-00107-3>
- Lockwood, E., Reed, Z., & Caughman, J. S. (2017). An analysis of statements of the multiplication principle in combinatorics, discrete, and finite mathematics textbooks. *International Journal of Research in Undergraduate Mathematics Education*, 3(3), 381–416. <https://doi.org/10.1007/s40753-016-0045-y>
- Lockwood, E., Wasserman, N. H., & Tillema, E. S. (2020). A case for combinatorics: a research commentary. *The Journal of Mathematical Behavior*, 59, 1–15. <https://doi.org/10.1016/j.jmathb.2020.100783>
- Morcos, J, & González-Martín, A. S. (in press). Permutations with indistinguishable objects: inconsistencies in the university mathematics training of prospective secondary teachers. Accepted for presentation in the *Fourteenth Congress of the European Society for Research in Mathematics Education (CERME14)*. ERME.
- Saldaña, J. (2013). *The coding manual for qualitative researchers* (2nd ed.). Sage.

PROSPECTIVE MATHEMATICS TEACHERS' PERCEPTIONS ON COLLABORATIVE LESSON STUDY

Galit Nagari-Haddif, Abraham Arcavi and Ronnie Karsenty

Weizmann Institute of Science, Israel

This study explores the perceptions of Israeli prospective mathematics teachers who participated in a Lesson Study process as part of their teacher education program. Conducted over the course of a year in a secondary school setting, the research involved 11 prospective teachers working in small groups to collaboratively plan and teach lessons. The findings highlight the contribution of Lesson Study in fostering collaborative skills, encouraging creativity, and promoting professional reflection, alongside challenges related to group dynamics and uneven participation. The findings suggest that Lesson Study may serve as an effective simulation for preparing prospective teachers for their future roles, equipping them with teamwork skills, openness to new ideas, and reflective practices.

INTRODUCTION

Teacher education programs worldwide increasingly emphasize the importance of equipping future educators not only with strong content knowledge but also with the skills necessary for continuous improvement of their teaching practices. One approach that has gained considerable attention for its potential to foster collaborative professional development is Lesson Study (LS). Originating in Japan, LS is a structured process that involves a cycle of collaborative lesson planning, observation of the lesson's implementation, analysis, and revision (e.g., Murata, 2011). This study explores the perceptions of prospective teachers (PSTs) regarding their participation in an LS process as part of their teacher education program. By considering both initial expectations of the PSTs before the process and their post-hoc reflections, the study provides a window into their experiences and allows for extracting insights regarding the contribution of LS as part of pre-service teacher education programs.

THEORETICAL BACKGROUND

The implementation of LS is collaborative in nature, as groups of teachers work together to plan, observe, and analyze lessons, thus refining instructional strategies through iterative cycles (e.g., Fujii, 2018). Research has shown that LS positively impacts teachers' pedagogical knowledge, self-efficacy, and student learning outcomes by fostering critical reflection on teaching practices and learning principles (Hiebert et al., 2003; Lewis et al., 2006). While LS is widely implemented among practicing teachers, its adaptation for PSTs, especially in mathematics and in non-Japanese contexts, remains underexplored. Studies suggest that LS can potentially encourage PSTs to learn collaboratively and develop their teaching skills through observation, reflection, and peer feedback (Hourigan & Leavy, 2019; Fernandez & Zilliox, 2011). In contrast to traditional settings where PSTs often plan lessons individually, the

collaborative nature of LS can contribute to professional growth by providing them with a structured, supportive peer-centered environment (Cajkler & Wood, 2014). Schipper et al. (2020) highlighted the potential of LS in helping PSTs bridge theoretical knowledge and practical teaching and in fostering meaningful peer interactions. This study explores whether such potential is realized by examining Israeli PSTs' experiences with LS. The research question we posed was: What are PSTs' perceptions of their participation in an LS process during their practical training?

METHODOLOGY

Setting (1): Program and participants. The study took place in a teacher education college for women. Eleven PSTs, in their second year of studies, participated in a year-long practical training at a secondary school (middle and high school levels). The PSTs were divided into four subgroups of 2-3 students (henceforth Groups A, B, C, D). In the first semester, each subgroup participated in lesson observation and analysis, engaged in individual lesson planning, and collaborated as a group to deliver substitute teaching. Concurrently, they were introduced to the LS model and observed recorded Japanese and Israeli lessons planned according to LS principles, in a teaching methods course taught by the first author of this paper, who also served as their pedagogical mentor throughout the training. In the second semester, the PSTs were encouraged, but not required, to conduct LS cycles. Groups A, B, and C collaboratively planned lessons (see details below), which were implemented in the school by group representatives, allowing the group members to observe the outcomes of their planning and teaching. Group D, however, opted for an alternative approach, with each member independently planning their own lesson.

Setting (2): Description of the lessons planned by the groups. Groups A and B developed highly detailed and meticulously prepared lessons, while Group C's lesson was less refined. Group A, as pioneers in the process, showcased their planning stages, challenges, and decisions, serving as a model for other groups. Their 7th-grade lesson on identifying and labeling angles involved experiential activities such as drawing and using rubber bands on pegboards (Figure 1a). In the implementation of this lesson students seemed engaged, although some difficulties arose with correctly labeling angles and handling the physical artifacts. Inspired by Group A, Group B created a lesson on the circumference and area of a circle for an advanced-level 7th-grade class. The lesson plan employed interactive activities to guide students in discovering the relationship between circumference and diameter, introducing pi (π), and used visual methods like rearranging circle wedges to approximate a rectangle and derive the formula for the area of a circle. In implementing this lesson, the PST who taught it encountered challenges with classroom management and student instructions, but managed to keep the exploratory and interactive nature of the plan. Group C, influenced by Group A's emphasis on experiential learning, designed a lesson on probability. Beginning with games involving dice and cards, they introduced fairness and other probability concepts using objects such as coins and 4-sided spinning tops

(Figure 1b). In the implemented lesson, these activities took less time than what the group has planned, and the teacher improvised additional activities.

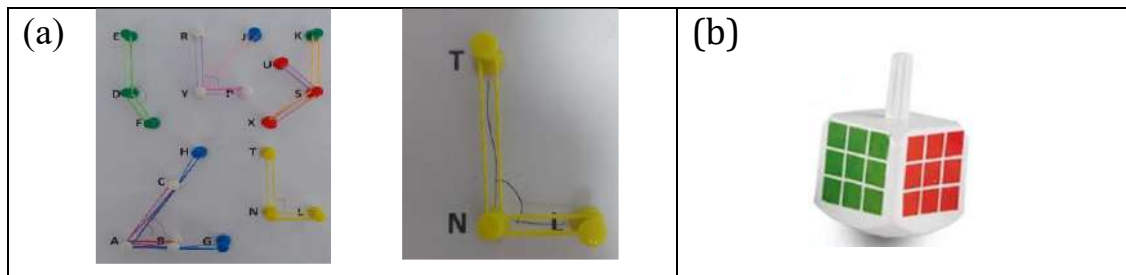


Figure 1: (a) Rubber band and pegboard artifacts (b) Four-sided spinning top

Data collection. Each PST participated in two semi-structured individual interviews conducted at the end of the first and the second semesters (henceforth: Int1, Int2). All interviews were transcribed. The PSTs also submitted an individual written reflection at the end of each semester (henceforth: RefA, RefB). The interviews and reflections covered various aspects of the PSTs' practical training, including their views on LS before the process (e.g., "*What are your thoughts on implementing LS?*") and afterward (e.g., "*Describe your experience of implementing LS*").

Data analysis. We employed a bottom-up approach to explore the PSTs' perceptions of LS before and after experiencing it, identifying key themes that emerged from the data, such as the PSTs' views on collaborative work, their observations of student reactions, and their reflections on the overall LS process, both within their groups and in relation to the other LS sources they were exposed to in the course.

FINDINGS

In this section we present and illustrate themes found in the analysis. Notations such as Stu1(A), Stu2(C) refer to different PSTs and the subgroup they worked in.

PSTs' perceptions before the lesson study experience

Perceived affordances of collaborative planning. Several PSTs noted the possible gains of working collaboratively, emphasizing that it can enrich and allow diverse ideas to emerge. Illustrative citations are: "working together can be really good" (Stu1(A), Int1); "it's an advantage to have multiple people thinking together" (Stu2(A), Int1). The group work was seen as a source for reassurance about the plan (e.g., "it gives confidence when there is agreement", Stu1(D), Int1).

Compromising conflicting views. The PSTs expressed concerns about potential clashes in beliefs or approaches when co-planning a lesson: "I imagine there will always be disagreements" (Stu1(A), Int1); "What if one person believes in one thing and another person in something else?" (Stu2(B), Int1). Some PSTs referred to situations where compromise might be needed, but difficult to achieve. For example:

I won't produce something that isn't 100% in my view [...] I have my opinion, and I express it. I am happy to hear a different opinion, but if it's something I strongly disagree with, it probably won't happen. (Stu3(A), Int1)

Of course, on very critical principles, there's no room for compromise. (Stu1(A), Int1)

I am very possessive about my way... I want it to be the way I envision it. (Stu1(C), Int1)

However, others were optimistic about the prospect of reaching agreements within the group, emphasizing the importance of flexibility and dialogue:

We can always find a compromise [...] It's not a life-changing matter here. (Stu2(B) Int1)

You need to learn to be flexible to work together [...] It won't always be the lesson I imagined. (Stu2(A), Int1)

If there's disagreement, we'll debate and resolve it together. (Stu1(D), Int1)

Implementing LS as weighing down the already demanding program. PSTs were concerned that the LS process would encumber on their existing resources, in terms of both the time needed ("sometimes the workload is too heavy", Stu2(B), Int1) or the mental effort required ("if it takes so much mental energy, it's not worth it", Stu3(A), Int1).

PSTs' perceptions after the lesson study experience

Gains afforded by the collaborative work. The PSTs noted various kinds of gains they saw in the collaborative LS process. For Example, Stu2(B) described her collaboration with Stu3(B) as follows:

Stu3(B) handled all the worksheets. From the beginning, she stated what she's good at and what she isn't, and I really appreciated that. She took responsibility for the technical aspects - designing worksheets, creating tables, and more. Writing was more difficult for her, so she asked us to do that part together. I felt the division of tasks was very fair. Stu3(B) is excellent in technical matters, which I'm not, and I'm strong in writing, which she's less comfortable with. We all contributed ideas, phrased them, and wrote them together. (Stu2(B), Int2)

The division of labor as described by Stu2(B) was not only efficient but also fostered mutual respect and appreciation. Other gains mentioned were the learning from each other's perspectives, the advantages of brainstorming, and the depth of detail reached:

The three of us have different ways of thinking and teaching approaches. It was interesting to see how each of us contributed our personal perspective to the group, adding uniqueness while also blending together. (Stu1(B), Int2)

Working together allowed us to think collectively and refine ideas through discussion and debate, even if it was sometimes confusing. (Stu2(A), RefB)

It was hard to schedule and reach decisions, but working together helped resolve dilemmas faster than planning alone. (Stu3(C), Int2)

It made us consider details we wouldn't have addressed otherwise. (Stu2(B), RefB)

Drawbacks and challenges of group work. In retrospect, some PSTs highlighted significant challenges involved in the need to collaborate, mostly around unequal participation and the tendency for some group members to become complacent:

It's easy to pass the work to someone else. If I don't do it, someone else will. (Stu1(A), Int2)

Sometimes, I would propose an idea, and the response was just, 'Great, let's move on,' without much further discussion. [...] I pushed them to join Zoom meetings and give input. I tried to give them space, but they didn't take it. (Stu3(A), Int2)

Working with Stu1(B) was hard. She assumed Stu2(B) and I would teach and she distanced herself, which frustrated us. We eventually asked her to contribute, and it improved (Stu3(B), Int2).

The need to navigate differing opinions and to compromise, which the PSTs worried about before they began the process, was still seen as challenging, for example:

In group work, you have to consider everyone's views, even if they don't align with your own. Sometimes it felt like someone had to give up on their idea for the group to move forward. (Stu2(B), RefB)

It's hard to merge everyone's approaches into a cohesive lesson. Even when teachers share similar beliefs, deciding what to do is difficult, and when beliefs differ, it's even harder. (Stu3(B), Int2).

Personal professional growth through LS. Various kinds of individual professional gains were mentioned by the PSTs. The quotes below illustrate some of these:

My creativity is relatively limited, and I'm concerned that creative activities might not be educational enough. Stu3(A) is very creative and I'm very realistic, so I think it was a good combination. I learned that things I wouldn't have done alone, like visual activities with boards and pins, can be both engaging and educational. It showed me that thinking outside the box can still be effective because the students learned a lot. (Stu1(A), Int2)

This citation depicts how the co-planning of the lesson, that involved concrete artifacts for learning about angles (see Figure 1a), expanded the spectrum of what the PST saw as "educational" activities and helped her accept creative instruction as effective. A similar shift in viewpoint was experienced by one of the Group B participants, regarding the development of formulas:

As a student, I preferred being given the formula and using it like a parrot. But LS made me reconsider, as understanding the origins of formulas had an added value for me. However, I still question its necessity for seventh graders - it might benefit only a few in the long term. Still, this experience probed my perspective and showed me the importance of fostering deeper understanding when possible. (Stu3(B), Int2)

This PST's reflection shows the complex form that "change in perspective" might take. In what seems a candid self-inspection, she acknowledges that the LS process exposed her to very different considerations than those she had as a student, yet that this new 'reconsidering' does not necessarily eliminate doubts. In the interview, the same PST shared how discussions with her LS peer also altered her look at classroom occurrences: "Stu1(B) emphasized the importance of student involvement, which made me rethink classroom dynamics and notice when some students weren't participating".

Another individual gain mentioned is linked to the aforementioned challenge involved in collaborating with peers:

This year, largely thanks to the lesson study, I learned the importance of patience and teamwork, not just in how I deliver lessons and material but in how I work with others. I realized that throughout my career, I'll encounter many situations where I'll need to collaborate with students or colleagues. This year gave me significant tools to approach such situations with patience and understanding. (Stu2(B), RefB)

Here, the personal "takeaway" relates to an affective aspect of the LS process.

Learning from other groups. PSTs reflected on how they benefited from observing and hearing about lessons of other groups. The citations below relate to Group A's lesson:

Their lesson on angles was so detailed and creative that it inspired me to aim for a similar level of planning in the future. (Stu1(C), Int2)

Seeing their collaboration motivated me to develop in similar ways and work more effectively in a team. (Stu3(C), Int2)

Their active, student-led lesson raised the bar for all of us. It showed the importance of engaging students in constructing their own understanding rather than simply delivering formulas. (Stu2(B), RefB)

Similarly, Group B's lesson was reflected upon:

They started with a dull idea but explored new ones, like approximating a circle's area through triangles—a method I hadn't seen before. It was fascinating to see how they let students discover the approximation themselves. (Stu3(A), Int2)

Students' reactions and ensuing adjustments. Several PSTs specifically noted the issue of how the lesson plan had anticipated (or not) students' actual responses, with some PSTs perceiving unforeseen students' reactions as a shortcoming of the plan, while others seeing it as unavoidable and even as an opportunity to improvise:

The task definition wasn't clear enough. We needed to refine our lesson and better anticipate student responses. (Stu3(B), RefB)

The process paid off because the lesson turned out amazing. It was interesting to see the students' reactions and engagement [...] No matter how well we prepare, unexpected moments will arise. I didn't anticipate all the possible student questions. (Stu1(B), Int2)

We assumed the students would struggle with probability, but they understood much faster than expected. This required improvisation during the lesson, which the collaborative planning helped prepare me for. (Stu2(C), Int2)

We started defining angles, only to realize the students already knew terms like acute, obtuse, and straight angles. Despite this, the lesson was memorable and enjoyable for them. (Stu3(A), Int2)

Future ideas and suggested variations: Participants offered ideas for refining the LS process to balance collaboration and individual teaching needs, for example: "Start with group brainstorming to generate ideas and frameworks, then let each teacher adapt

the lesson to their classroom. This combination maximizes creativity while respecting individual teaching styles and classroom needs" (Stu3(A), Int2). It was also proposed to involve experienced teachers in the LS process: "Veteran teachers know what works and can ground ambitious ideas coming from newer teachers" (Stu1(A), Int2).

To sum up the findings: Prior to the LS implementation, the PSTs could see value in collaborative lesson planning, but were worried about needing to compromise different views and about the extra-resources to be invested. After their LS experience, PSTs still pointed to some challenges involved in collaborating with each other, but appreciated the various affordances of the LS model, in terms of mutual learning, personal growth and broadening of perspectives.

DISCUSSION

Although LS is traditionally designed for practicing teachers, as it relies on their experience and aims to address immediate classroom needs, this study highlights its potential value for PSTs and its unique characteristics in this context. In the study, we explored PSTs' perceptions about LS, prior to and after experiencing this model. One of the key findings is that PSTs' views on collaborative lesson planning evolved from concerns (mainly) around individuality and self-preferences, towards acknowledging the benefits of working with others and mutual learning (within and across groups). While the LS process still involved challenges, it appears that this experience brought PSTs closer to the real world of practice, awakened their sensitivity towards students' needs and highlighted the social nature of the teaching profession. Teamwork was emphasized, including the integration of diverse perspectives and the refinement of lesson designs through dialogue and joint decision-making. These findings align with Schipper et al.'s (2020) ideas about the potential of LS at the preservice level.

Although the study did not delve into individual trajectories of PSTs, the findings provide a window into some PSTs' evolving perspectives. For example, Stu3(A) was initially concerned with compromising her principles ("I won't produce something that isn't 100% in my view"), and was worried about investing too much "mental energy", yet she emerged as a leader and a creative teacher while learning to navigate frustrations regarding more passive peers. Her experience underscores the dynamic nature of collaboration and the importance of balancing individual contributions with collective goals, echoing Murata's (2011) findings. Her words at the end of the semester vividly illustrate the gains she saw in LS, and specifically in collaboration: "I'm glad I experienced both regular planning and LS. I hope to work with a [school] team that could plan lessons together like this (even if only partly, or just in outlines, and from there each teacher could take it and adapt it to her class)" (Stu3(A), RefB).

The study also revealed the potential of LS in encouraging PSTs to experiment with innovative teaching methods and to adapt their teaching strategies to diverse classroom contexts. In line with prior research (e.g., Hourigan & Leavy, 2019), we found that planning lessons within peer groups inspired PSTs to expand their repertoire of

approaches, deepen their understanding of student thinking and respond flexibly to unexpected classroom dynamics. In conclusion, this pilot study provides a “proof of concept” that LS can serve as an effective platform for preparing PSTs for the collaborative demands of their future careers. LS enables PSTs to explore the wholeness of the teaching profession - mathematics, didactics, planning, collaboration, and student-centered practices.

Acknowledgment

The study was supported by the Israel Science Foundation (Grant #1219/20) and the Weizmann Institute of Science. We thank the PSTs who agreed to take part in the study.

References

- Cajkler, W., & Wood, P. (2015). Lesson Study in initial teacher education. In P. Dudley (Ed.), *Lesson Study: professional learning for our time* (1st ed., pp. 107–127). Routledge. <https://doi.org/10.4324/9780203795538-5>
- Fernandez, M. L., & Zilliox, J. (2011). Investigating approaches to lesson study in prospective mathematics teacher education. In L. C. Hart, A. S. Alston, & A. Murata (Eds.), *Lesson study research and practice in mathematics education: Learning together* (pp. 85–101). Springer. https://doi.org/10.1007/978-90-481-9941-9_7
- Fujii, T. (2018). Lesson study and teaching mathematics through problem solving: the two wheels of a cart. In M. Quaresma, C. Winsløw, S. Clivaz, J. P. da Ponte, A. N. Shúilleabháin, & A. Takahashi (Eds.), *Mathematics lesson study around the world* (pp. 1–21). Springer. https://doi.org/10.1007/978-3-319-75696-7_1
- Hiebert, J., Morris, A. K., & Glass, B. (2003). Learning to learn to teach: An “experiment” model for teaching and teacher preparation in mathematics. *Journal of Mathematics Teacher Education*, 6, 201–222. <https://doi.org/10.1023/A:1025162108648>
- Hourigan, M., & Leavy, A. M. (2019). Learning from teaching: pre-service primary teachers’ perceived learning from engaging in formal Lesson Study. *Irish Educational Studies*, 38(3), 283–308. <https://doi.org/10.1080/03323315.2019.1613252>
- Lewis, C., Perry, R., & Murata, A. (2006). How should research contribute to instructional improvement? The case of Lesson Study. *Educational Researcher*, 35(3), 3–14. <https://doi.org/10.3102/0013189X035003003>
- Murata, A. (2011). Introduction: Conceptual overview of lesson study. In L. Hart, A. Alston, & A. Murata (Eds.), *Lesson study research and practice in mathematics education* (pp. 1–12). Springer. https://doi.org/10.1007/978-90-481-9941-9_1
- Schipper, T. M., Goei, S. L., Van Joolingen, W. R., Willemse, T. M., & Van Geffen, E. C. (2020). Lesson study in Dutch initial teacher education explored: its potential and pitfalls. *International Journal for Lesson & Learning Studies*, 9(4), 351–365. <https://doi.org/10.1108/IJLLS-04-2020-0018>

TRANSITION FROM DYNAMIC TO STATIC DIGITAL ARTIFACTS TO PROMOTE META-LEVEL LEARNING: A CASE STUDY WITH SCALES AND EARLY ALGEBRAIC DISCOURSE

Bernardo Nannini, Samuele Antonini

Department of Mathematics and Computer Science, University of Florence, Italy

This study explores the use of digital artifacts to support students' introduction to algebraic discourse. By adopting a discursive framework, we investigate narratives and routines emerging in middle school students' discourse who interact with digital artifacts representing scales in task situations requiring to determine the value of unknown weights. Activities were designed to move from dynamic scales that provide feedback to users' actions to static scales where feedback is inhibited. Analysis shows that dynamic scales fostered the emerging of a variety of different procedures, and the transition to static scales prompted students' discussion on the procedures themselves and their functioning, moving in this way their discourse on a meta-level.

INTRODUCTION AND THEORETICAL BACKGROUND

Research has widely documented the complexity of the teaching and learning of algebra and studied students' difficulties that include giving meaning to unknowns, equal sign as a relationship, and dealing with algebraic symbols and equations (e.g., Kieran, 1992). A stream of research focused on early algebra, on processes involved in algebraic thinking and on the study of effective approaches to introduce it to students (e.g., Carpenter et al., 2005; Kaput, 2008). With this background, in this paper we report on a study framed in the theory of *Commognition* (Sfard, 2008) focusing on a didactical approach with the use of digital artifacts for introducing students to *algebraic discourse*. This study is part of a wider research project on middle and high-school students' learning of school algebra with support of digital environments.

Commognitive approach

In this paper, in line with *Commognition* (Sfard, 2008), we see school algebra as a specific discourse with its own *keywords* (e.g., 'unknown', 'variable'), *visual mediators* (e.g. a symbolic equation), *endorsed narratives* (e.g., "1 is a solution for $2x+2=4$ ") and *routines* (e.g., solving procedure to determine the solution of a linear equation). In this perspective learning is seen as becoming progressively more able to participate in mathematical discourse and studying students' mathematical learning means to analyze their discourse and how it changes. In this regard, it is possible to differentiate between *object-level* and *meta-level* learning. The first one involves an expansion of a discourse which enriches with new keywords, visual mediators, narrative and routines. On the other hand, meta-level involves a change in the *metarules* of a discourse, which defines ways of participating in the discourse and

determining what can be endorsed in it. To investigate changes in students' discourse it can be useful to focus on their routines. Lavie et al. (2019) define routines as couples (*task, procedure*): a task is what a person interprets that needs to be obtained in a given *task situation* (a setting in which she considers herself bound to do something); a procedure is the prescription of actions that the person recognizes to be done to accomplish the task. Routines in students' discourse can be used as units of analysis to investigate mathematical learning (Lavie et al., 2019). For instance, they could grow in *flexibility*, when the same task is accomplished by different procedures, or in *applicability* when the range of tasks for which the same procedure is suitable enlarges. Moreover, a crucial aspect of routines involves their *substantiation*, which relates to how a student justifies the used procedure and its functioning to accomplish a task.

The project in which this study is situated has the goal of studying the role of students' participation in a discourse developed through activities with digital and interactive environments (Antonini et al., 2023; Baccaglini-Frank, 2021) in supporting their participation in algebraic discourse. In this paper, we focus on some activities with digital artifacts representing scales based on the balance model. This model is widely used in teaching linear equations, for example, to deal with unknowns or to introduce principles of equivalence for equations (for a review, see Otten et al., 2019).

Description of the designed artifact and its use for this study

The activities in focus in this study revolve around a digital artifact designed in Desmos (Figure 1) and presented to students on tablets. It consists of a two-pan scale with colored shapes realizing different weights. By dragging these weights on the screen, they can be put on or removed from the pans. The shapes are labelled with numbers corresponding to the value of the weight or can have unknown value. The scale gives feedback to user's actions by dangling to the right/left or staying balanced accordingly with the total value of the weights on its pans. A broader analysis of this artifact's affordances and their didactical potentialities is discussed in Weingarten et al. (2024).

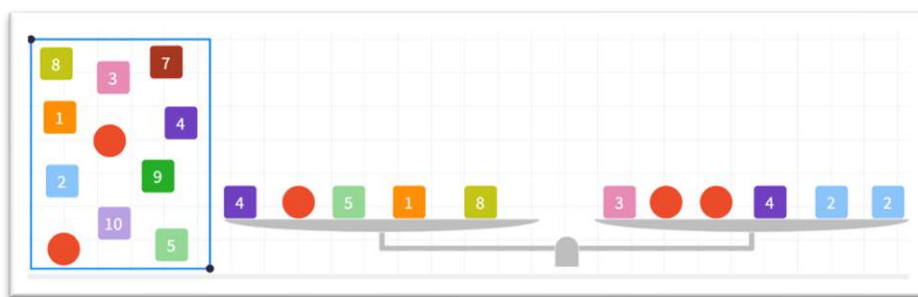


Figure 1: Screenshot of balanced scale with a box on the left where to put removed weights or from which to take new weights.

The didactic approach of the study in which this paper is embedded involves having students produce narratives about scales, such as to describe the conditions for a scale to be balanced/unbalanced, and routines to answer task situations with scales, such as to identify the value of unknown weights. Then, some of these elements of students'

discourse on scales could be used to move to an initial discourse about equations, in particular on principles of equivalence and solving procedures.

Research objective

With the goal of fostering this scale-equation transition, we decided to propose activities in which students interact with dynamic and interactive scales and other activities with pictures of scales, thus static and not providing feedback. Hereon we will refer to these two kinds of scales respectively as *dynamic* or *static*. The choice of using first dynamic and then static scales was made under the hypothesis that the inhibition of feedback on the static ones could prompt students' reflection on whether the procedures previously used with dynamic scales are still suitable in this second case (applicability) and this could promote the production of narratives about the procedures themselves and their functioning (substantiation). This paper intends to explore this hypothesis by analyzing routines and narratives emerging in students' discourse. Specifically, we investigate the following questions: *What procedures emerge in students' discourse when asked to determine the value of unknown weights in dynamic scales? What is the role of the transition to static scales on these procedures?*

METHODS

We conducted a case study with a class of 24 students of 8th grade from an Italian middle school. The class was novice to algebraic discourse, with no previous experience with equations and their resolution. The class participated in 4 weekly sessions of 90' each, taking place in a room equipped with cameras and microphones. Sessions were conducted by the first author of this paper (hereon 'tutor'). The mathematics teacher of the class was also present as an observer. For every given task situation, students worked in 8 groups of three, the same during the sessions, and then all groups were invited by the tutor to share their work with the others. Each group was provided with a tablet, paper and pens. The activities in focus in this study refer to the first two sessions. Collected data consist of the video recordings of the meetings, all written productions, and screen recordings of each group's tablet. In this paper we focus on three kinds of task situations. The first one (TS1) involves a dynamic scale (as in Figure 1) with the request: *The scale that you see is balanced. If possible put on [or take off] the pans 5 objects [or other number] so that the scale at the end will still be balanced.* The second kind of task situation (TS2) requires determining the value of an unknown weight on a dynamic scale. The request is the following: *The scale that you see is balanced. Could you discover the weight of the red ball?* Finally, the third kind (TS3) has the same request as the second one but on a static scale with a writing on top: *This is a picture, thus nothing can be moved, but you can write on it if it helps.*

SUMMARY OF THE CASE ANALYSIS

In what follows we provide four snapshots to report on some main aspects that emerged from the case study analysis. Names are pseudonyms. A general discussion of these snapshots in relation to the research objective is presented in the next section.

Snapshot 1 – At the end of TS1

At the end of the first activities with dynamic scales, different narratives emerged in the discussion (e.g., “[the scale is balanced] when the weight is equal on both sides”, “when the sum of the weights is the same”, or “[if adding or removing weights, the scale remains balanced] when the sums of those weights are the same”). At the end, these narratives were synthesized on the whiteboard by the tutor in the following two: (1) “A scale is balanced when the sum of the weights on the left pan is equal to the sum of the weights on the right pan”; (2) “If a scale is balanced and we add or remove the same sum of weights on the left and on the right, the scale remains balanced”.

Snapshot 2 – Different procedures for TS2

Afterward, groups were involved with activities of the kind TS2, requiring determining the value of an unknown weight in some given dynamic scales. Three different procedures emerged in the classroom (see Table 1).

Procedure	Groups using it	Written description after class discussion
<i>Weighing</i>	Groups 1-8	“Strategy 1. I leave one ball on one side and on the other I make some attempts until the scale is balanced”
<i>Reverse-Weighing</i>	Groups 1, 2, 5, 8	“Strategy 2. We take off one ball from one side and on the other we take off weights until the scale is balanced”
<i>Removing</i>	Groups 2, 8	“Strategy 3. I take off from both pans the same weight until on one pan only one ball remains and on the other pans only numbers”

Table 1: Description produced in the class discussion of the procedures used in TS2.

The first one, used by all groups, consisted in emptying completely the scale except for one unknown weight on one pan and then putting known weights on the other until the scale resulted balanced. Note that this procedure exactly matches how a two-pan scale can be used to weigh objects. For brevity, we name it the *Weighing-Procedure*. A second procedure, used by four groups and quite similar to the previous one, consisted in removing one unknown weight from one pan and then known weights from the other until the scale was balanced. If a balanced situation is reached one can conclude that the same amount of weight has been removed from both sides and thus determine the unknown weight. We name it the *Reverse-Weighing-Procedure*. Lastly, two groups also used a third different procedure consisting in removing equal weights, or a combination of the same total value, from both pans until reaching a situation in which the scale resulted balanced, and the unknown weight is identified (for instance if the scale has only an unknown weight on one side and known weights on the other). We name it the *Removing-Procedure*. We observe that, in the eyes of an expert, this last procedure recalls that for solving linear equations consisting in cancelling out numbers and unknowns in both sides to obtain an equivalent equation. Differently, the other

two, whose functioning is based on the scale's feedback, do not find correspondence in equations. At the end of the class discussion, a description for each procedure (called "strategies" by the students) was reached and written on the board (see Table 1).

Snapshot 3 – Episodes from TS3

The following activities again required determining the value of an unknown weight, but this time with a static scale (TS3). We present two brief episodes from two groups involved with the first activity of this kind based on the scale shown in Figure 2a, in which a purple square has an unknown weight.

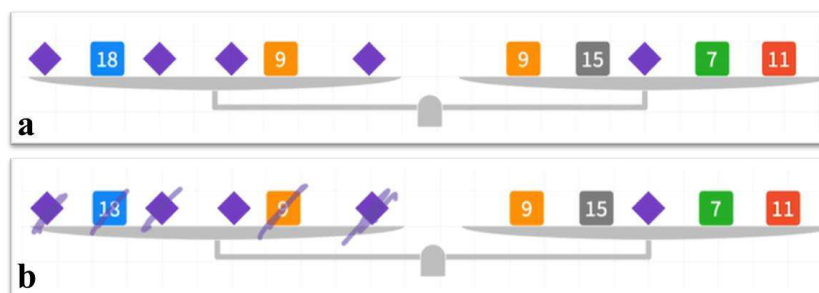


Figure 2: Static scale in first activity of TS3 (a), with inscriptions by Group 5 (b).

Group 5 – Edo and Mat

Right after reading the request, the students try to drag some weights, realizing that it is not possible. After two minutes of silence Edo intervenes:

Edo: If the scale would move, we could figure it out, but without how can you do it? We cannot [...] Because you could take off here, you could leave as before one purple square and, on the other, you add some of these numbers and then you see. But the scale doesn't move. So, I think it is impossible.

Then Mat draws some inscriptions on the scale following Edo's previous description (Figure 2b). After this, the students stop without reaching a solution.

Group 8 – Ale and Fra

Ale tries to drag away an 11-weight and sees that it is not possible. Then he silently makes some round marks on the picture: first, on an 11- and on a 7-weight on the right and on an 18-weight on the left; then on one purple square on the left and on one on the right; finally on a 9-weight on the right and on one on the left. Then he concludes:

Ale: Fifteen divided by three equals five. So, this thing here is equal to five.

Fra: But are you sure?

Ale: Yes, because there are three [purple squares] left and here there is fifteen and it is balanced.

Fra: Yes, but if it is blocked, you don't know if the weight is the same!

Ale: Yes, because... eleven and seven is eighteen and here there is eighteen, so it is like taking off eighteen here and here. Then there is one little square, and I took off one little square here. Nine and nine. So, it remains fifteen

and three of these, and it should be balanced. And then one can just do fifteen divided by three.

These two short episodes brought an example of how the groups differently tried to reproduce previously effective procedures in this new task situation. Edo from group 5 attempted to reconstruct the so-called Strategy 1 (“as before”) and following him Mat also produced some visual mediators to represent the corresponding actions (Figure 2b). However, the absence of feedback from the scale, crucial to the functioning of this strategy, undermines their attempt. Differently, in Group 8, Ale reconstructed Strategy 3 on the picture with some cancellation marks. Interestingly, Fra was initially not convinced of the applicability of Ale’s procedure, observing that “you don’t know” whether the (static) scale is balanced. Ale made explicit why we know that at the end the scale “should be balanced” by describing step by steps his procedure. Note the conditional in this last sentence which seems to stress the fact that the scale did not actually move but its behavior can be predicted.

In the comparison of these two episodes and in the brief dialogue between Ale and Fra, a substantial difference between the two procedures begins to emerge in the students’ discourse: whereas the former requires what we “see” (feedback from the scale), the applicability of the latter procedure is assured by what we “know” (about our actions and general properties of scales). As we shall see, this aspect was central in the class discussion after the TS3 activities and which we briefly present in the next snapshot.

Snapshot 4 – Discussion at the end of TS3

The tutor opens a discussion with a question about the applicability with the static scales of the procedures used before with the dynamic scales. Students agreed that the strategies 1 and 2 cannot work. They justified this answer by referring to the absence of feedback: e.g., “Because we cannot know when it will be balanced” or “Because the balance would always stay the same anyway [...] even if you remove like 15 or 9, it stays the same”. Then the tutor asks if strategy 3 could work. We show the discussion involving Tom (Group 3), Anna (Group 2) and Leo (Group 6).

- Tom: No [it cannot work]. For the same reason as before. One can think of removing them [the weights], but the scale doesn’t change. [...]
- Anna: Yes [it works]. Because it says that one takes off from both pans the same weight and we can know this thing. For example, if I take off nine and nine, I know it is the same weight... and therefore one can know! [...] One can know of having taken off the same weight and therefore that the scale is still balanced, if we take off the same weight. [...]
- Leo: It depends... One must always take off the same quantity, otherwise... If one always takes off the same quantity, then one can imagine that it always remains balanced, but if one takes off random numbers, let’s say, then no. [...] Because being balanced at the beginning, if one takes off the same quantity anyway it remains balanced... because if the sum is equal at the beginning, after taking off the same sum, it is necessarily equal even later.

In this short excerpt we can observe how the students' discourse is shifted to a meta-level, where the objects of the discourse are the procedures, their applicability and the possible actions that a person might take. While Tom refers only to the lack of feedback of the scale, in Anna's argumentation, the verb 'to know' is used several times in relation to the fact that the scale would remain balanced. In her view, this act of knowing is independent of the scale's feedback but depends on the subject's own action ("if I take off nine and nine, I know it is the same weight!"). Leo further emphasizes this fact by explicating two kinds of actions when there is no feedback (taking off weights at "random" or "the same quantity") and that only in the second case, which corresponds to strategy 3, "one can imagine that it always remains balanced". The conclusion of the episode presents further interesting elements. To justify why the scale remains balanced, Leo produces a general argument (in the last utterance there are no references to the scale but to the equality of sums) which is based precisely on those endorsed narratives that were previously produced by the class (snapshot 1).

DISCUSSION AND CONCLUDING REMARKS

The analysis conducted highlighted an expansion of students' discourse in response to TS1 and TS2. Specifically, their discourse enriched with endorsed narratives describing general properties of the scales (snapshot 1) and with different routines involving procedures for determining unknown weights (snapshot 2). In response to the first research question, the analysis revealed three main procedures that emerged in the class (Table 1). Some groups used more than one procedure for the same task, an aspect that highlights some flexibility in these routines. Addressing the second research question, the analysis showed how in the transition from dynamic to static scales, with the inhibition of the dynamicity and of feedback, the tutor's interventions and the dialogue among students' different points of view, promoted a crucial discussion on the applicability of the previous procedures. Whereas in TS2 it was possible to perform actions and observe the feedback from the scale, in TS3 it became decisive to imagine these actions and the scale's corresponding feedback. The absence of dynamicity fostered students' production of written visual mediators to talk about the steps of these procedures, an aspect that, as suggested by Antonini et al. (2023), can play an important role in the transition to algebraic discourse. On the other hand, the absence of feedback promoted a discourse characterized by a tension between empirical observations (e.g., "and then you see" by Edo in snapshot 3 or "the scale doesn't change" by Tom in snapshot 4) and endorsed narratives (e.g., "one can know" by Anna, followed by Leo's general justification in snapshot 4). This tension moved the students' discourse on a meta-level, whose objects were the procedures themselves and their substantiation. This analysis contributes to what showed by Baccaglini-Frank (2021) on digital interactive artifacts that can support students' participation in mathematical discourse by providing "protagonists" for meaningful narratives. Indeed, our study shows that carefully designed transitions from dynamic to static artifacts, in which feedback is inhibited, could support students' production of narratives to justify these protagonists' behavior and students' actions on them.

Finally, the development of students' discourse led them to differentiate the *Removing-Procedure* from the other two. This aspect is significant in at least two respects. Firstly, it shows how the *Removing-Procedure* and its substantiation can be promoted. This was one of the teaching objectives of the activities due to its link with the solving procedure for linear equations based on the principles of equivalence and since students' substantiation of the *Removing-Procedure* could support them in substantiating the procedure for equations, an aspect often problematic for students (Kieran, 1992). Secondly, the tension between empirical observations and production of narratives initiated a process of changing students' metarules for endorsing narratives (we can "know" that something is true even if we do not "see" it). In other terms, a meta-level learning was promoted and this may become an initial seed for a future algebraic discourse on equations based on theoretical principles.

Acknowledgement

Study part of the project DynaMat (PRIN2020BKWEXR) and founded by Fondazione CARIPT (Ricerca e Innovazione 2024 – n. 32678). Experimental work conducted at CARME (www.carme.center), UNISER Pistoia Srl, Italy.

References

- Antonini, S., Bonadiman, C., & Nannini, B. (2023). Dynamic interactive mediators in discourse on indeterminate quantities: A case study. In M. Ayalon, B. Koichu, R. Leikin, L. Rubel, & M. Tabach (Eds.), *Proc. 46th Conf. of the Int. Group for the Psychology of Mathematics Education* (vol. 2, pp. 43–50). PME
- Baccaglini-Frank, A. (2021). To tell a story, you need a protagonist: how dynamic interactive mediators can fulfill this role and foster explorative participation to mathematical discourse. *Educational Studies in Mathematics*, 106(2), 291–312.
- Carpenter, T.P., Levi, L., Franke, M.L., & Koehler Zeringue, J. (2005). Algebra in elementary school: Developing relational thinking. *ZDM Mathematics Education*, 37(1), 53–59.
- Kaput, J.J., Carraher, D.W., & Blanton, M.L. (2008). *Algebra in the early grades*. Routledge.
- Lavie, I., Steiner, A., & Sfard, A. (2019). Routines we live by: From ritual to exploration. *Educational Studies in Mathematics*, 101(2), 153–176.
- Kieran, C. (1992). The learning and teaching of school algebra. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning: A project of the National Council of Teachers of Mathematics* (pp. 390–419). Macmillan Publishing Co, Inc.
- Otten, M., Van den Heuvel-Panhuizen, M., & Veldhuis, M. (2019). The balance model for teaching linear equations: a systematic literature review. *International Journal of STEM Education*, 6(1), 1–21.
- Sfard, A. (2008). *Thinking as communicating: Human development, the growth of discourses, and mathematizing*. Cambridge, UK: Cambridge University Press.
- Weingarden, M., Lisarelli, G. & Baccaglini-Frank, A. (2024). Exploring Mathematical Learning Opportunities Afforded by a Balanced Scale Digital Activity. *Digital Experiences in Mathematics Education*, 10, 382–394.

“HOMEWORK SHOULD ONLY PREPARE FOR THE EXAM!” – STUDENTS’ ASSESSMENT OF (POTENTIAL) OBLIGATORY HOMEWORK IN MATHEMATICS EDUCATION COURSES AT UNIVERSITY

Silke Neuhaus-Eckhardt, Hans-Stefan Siller

Julius-Maximilians-University Würzburg

Homework at university provides out-of-class learning opportunities, but many students do not use those appropriately, especially if the assignments are voluntary. To gain insight into the question of whether homework should be obligatory, we analysed the conversion of a mathematics education course from voluntary to obligatory homework. For this contribution we used qualitative content analysis to analyse the students’ answers on how they rate (potential) obligatory homework. Students in both courses rated an obligation as rather negative and reasoned with additional stress and references to other courses. In addition, students with obligatory homework saw low relevance in their tasks and many of them demanded homework to (only) prepare them for their exam.

INTRODUCTION

Lecturers often provide additional homework for students to do after class, sometimes with the opportunity to submit their solutions for feedback, because they think that only listening to a lecture does not provide enough learning opportunities (Koban et al., 2020, Pritchard, 2015). However, there are only a few empirical studies analysing this relation for mathematics and even less for mathematics education courses and their results are inconsistent (e.g., Koban et al., 2020). But as many students do not do voluntary assignments, some lecturers would prefer obligatory homework in the hope that this obligation will motivate their students to benefit from this learning opportunity (Halcrow & Dunningan, 2012). But even an obligation does not seem to be advantageous for all students at least regarding performances in exams (Koban et al., 2020, Krapf & Liebendörfer, 2021). One reason may be that students tend to engage in inappropriate learning behaviours if they are overwhelmed or view obligatory homework only as a demand not a learning opportunity. As such, to answer the question of whether homework should be obligatory or voluntary at university, the students’ assessment of (obligatory) homework as a helpful learning opportunity is important. If most students reject an obligation, have no time to do homework or see no relevance in it, it can be assumed that they will not make appropriate use of the provided learning opportunity and that the increased effort required by lecturers to develop, provide, and correct homework may be pointless. We analysed the conversion of a course in mathematics education from voluntary homework in 2022 to obligatory

homework in 2023. For this contribution we concentrate on the question how students perceive (obligatory) homework in their course as a (helpful) learning opportunity.

HOMEWORK AT UNIVERSITY

Homework at university can be defined as tasks assigned to students by their lecturers that should be carried out after class, which is analogous to the definition of homework at school (Cooper, 1989). It is often assumed, that doing such additional exercises benefits the students' learning outcomes, but there are only a few empirical studies analysing this relation. Those studies mostly analyse mathematics not mathematics education courses and show inconsistent results (e.g., Koban et al., 2020). For example, Ufer (2015) identified that students doing voluntary homework had better grades in the final exam, whereas Halcrow and Dunnigan (2012) could not identify a clear trend.

Despite the lack of clarity about the relation between homework and study success, some lecturers demand homework at university to be obligatory because they think that many students do not work on voluntary assignments (e.g., Halcrow & Dunnigan, 2012, Lunsford & Pendergrass, 2015). But students may also not use homework adequately as a learning opportunity if they are forced to hand in their solutions. Again, there are only a few studies that have analysed the usefulness of obligatory homework so far. In the study of Koban et al. (2020, introductory mathematics), grading daily homework led to a higher performance on homework but only medium-ability students, who did homework, showed a significantly higher exam performance. Krapf and Liebendörfer (2021) analysed the conversion of a course in elementary mathematics from voluntary to obligatory homework. There was a higher participation and pass rate in the exam compared to the previous semester, but not a better average grade. In an additional questionnaire about how the obligation had an impact on their learning, the students in the study of Krapf and Liebendörfer (2021) gave mainly positive arguments, for example dealing more with the course's content or a feeling to be better prepared for the exam. But the students also mentioned more stress or excessive demands due to the obligation. Receiving feedback through correction, completing fewer assignments (because they divided them up within the learning group) or copying solutions, were only rarely mentioned.

First own results (Neuhaus-Eckhardt & Siller, under review) showed that students in courses with obligatory homework assess this obligation rather positive whereas students in courses with voluntary homework assess a potential obligation of their homework rather negative. This could indicate that students need to experience an obligation to assess it as positive for their own learning. In addition, the students' reasons how and why they view (obligatory) homework (not) as a learning opportunity can give us a deeper insight. For this, situated expectancy-value theory (Eccles & Wigfield, 2020) can be helpful.

SITUATED EXPECTANCY-VALUE THEORY

According to Eccles and Wigfield's situated expectancy-value theory (2020), students' performance related decisions, like doing obligatory homework or putting effort into it, can be explained by their expectancies for success or their task values.

Expectancies for success are defined as students' beliefs about how well they will perform on an upcoming task. Subjective task values are divided into intrinsic value (interest or enjoyment regarding a task), utility value (how well a student thinks this task fits in his or her present or future plans), attainment value (the relative personal/identity-based importance) and costs (negative aspects of a task).

Regarding homework at university, it is already shown, that students do not see a benefit in homework, if they are not graded or if they cannot get useful feedback (Halcrow & Dunningan, 2012). When we asked students in different undergraduate mathematics or mathematics education courses to give reasons for or against doing homework (Neuhaus-Eckhardt & Siller, 2024, under review), most of the students referred to utility values, especially preparing for the final exam, and therefore the students mostly showed an extrinsic motivation to do homework. According to the self-determination-theory (Ryan & Deci, 2000) extrinsic motivation does not lead to in-depth learning so that rather low learning outcomes are to be expected. Furthermore, if most students believe that the main purpose of homework is to prepare them for their exam, then they may assume that there is no benefit, only costs, in other homework tasks. But the tasks used in homework and written exams differ simply because of various external characteristics, for example time restriction and being able to use additional information. In addition, homework may thematize content, which is not needed for the exam but later in the study program or in a profession like a teacher. If students do not see an additional value, they may not put effort into tasks they view unhelpful to prepare them for their exam. Especially if those tasks are obligatory and the students need to gain enough points in a short time or feel overwhelmed, they may just copy solutions from other students or the internet (Liebendörfer & Göller, 2016; Robinson & Kuin, 1999).

Overall, it seems that dealing with (obligatory) homework could be beneficial for at least some students, but more studies, especially regarding mathematics education courses, are needed. In addition, possible positive relations may be influenced by how students assess their homework as a learning opportunity. When discussing making a course's homework obligatory, it is therefore important to analyse the students' point of view regarding (obligatory) homework in this course. An obligation may force the students to engage with the tasks and therefore their opinion about (obligatory) homework may be positive if they found the tasks helpful in retrospect. However, it may also be negative if they realise that not all tasks prepare them for the exam or if they have no capacity to do homework in a meaningful way.

RESEARCH QUESTIONS

We conducted a study to analyse the conversion of a course in mathematics education from voluntary to obligatory homework. For this contribution we wanted to get an insight into how the students in the course with voluntary homework and a year later the students in the course with obligatory homework view (potential) obligatory homework as a learning opportunity. As such, we focus on the following research question:

How do students with voluntary or obligatory homework rate (potential) obligatory homework for their mathematics education course and what reasons do they give for their assessment?

METHOD

The mathematics education course “Didactics of Analysis” for higher secondary school teachers (German Gymnasium, mostly third semester) was converted from voluntary homework (2022) to obligatory homework in the next year (2023). The course covers content like learners’ misconception or teaching strategies for analysis in school. The weekly homework in 2022 ($N_1 = 19$, 52.6% female) was voluntary and only very few students did the homework assignments which could be handed in for feedback. In 2023 ($N_2 = 31$, 35.7% female, 2.4% diverse) the students had to gain up to 60% of points in the weekly homework to get the admission to the written exam. Both times the students got exercises which were designed to deepen the students’ understanding of the content and to address the use of digital tools in analysis lessons at school.

We conducted a paper-based questionnaire (online in 2023) including an item asking the students on a 6-point Likert-scale from 1 (totally disagree) to 6 (totally agree) how useful they consider obligatory homework for their course. In addition, the students answered an open item asking why they think obligatory homework in this course to be (not) useful. In 2022 the questionnaire took place a few weeks before the final exam whereas in 2023 it had to be conducted after the final exam.

The students’ answers were coded using qualitative content analyses (Mayring, 2022) with a category system with 20 categories which was already developed and tested in an earlier study (Neuhaus-Eckhardt & Siller, under review). The data for 2022 was coded by a second coder and for 2023 it was coded a second time by the same rater after three weeks (approx. 80% compliance). Positive arguments for obligatory homework are for example to better understand the course’s content or preparation for exams. Examples of negative arguments are higher stress due to obligation or a lack of autonomy. However, many of the categories contain both negative as well as positively connoted arguments. For example, while some students stated that the quality of solutions would improve if homework was obligatory, others indicated the exact opposite.

RESULTS

As can be seen in table 1, the students of both courses rather disagreed to obligatory homework, which is also indicated by the small number of arguments in favour of obligatory homework.

Course	<i>M (SD)</i> (<i>Max</i> = 6)	Arguments for obligatory homework	Arguments against obligatory homework
Didactics of Analysis 2022 (voluntary, <i>N</i> = 19)	2.58 (1.71)	12	40
Didactics of Analysis 2023 (obligatory, <i>N</i> = 31)	2.26 (1.46)	9	81

Table 1: Assessment (on a 6-point Likert-scale from 1 “totally disagree” to 6 “totally agree”) and number of arguments for and against obligatory homework in the respective course.

In 2022 (voluntary homework) ten students mentioned that homework in this course is too time consuming, often referring to other courses (*n* = 7).

DA22_5 (translated): “In addition to Didactics of Analysis, there are several other courses that have obligatory homework exercises, which are already very time-consuming. It is not always possible for me personally to hand in additional exercises in Didactics of Analysis, which take even more time.”

Another argument often mentioned was that they would mind the lack of autonomy (*n* = 6) if homework was obligatory, for example:

DA22_15: “I don't think [an obligation] makes sense because I think that every student should decide for themselves which tasks and how much of the homework sheets they do. The homework sheets are useful and helpful for understanding, but sometimes you don't have enough time to do all the exercises because you also have other courses.”

But some students also gave arguments for an obligation. For example, they mentioned a better working attitude (*n* = 2) or relevance (*n* = 3) for their later life as a teacher. In addition, two students stated that adding an obligation to the curriculum would increase the value of this (in their opinion very important) course in contrast to other (less important) courses.

DA22_05: “I think an obligation makes a lot of sense, as the didactics lectures are the only professionally relevant lectures in the teacher training program. That's why they should be given more attention in the degree program (also in the curriculum).”

There are even fewer arguments for an obligation stated by students in 2023 (obligatory homework). Many of the students saw no relevance in the assignments (*n* = 18) and pointed out that the homework tasks were not useful for preparing for the exam (*n* = 9) which was also mentioned by three students in 2022. But only in 2023

students stated that the purpose of homework is only to prepare them for their exam. In addition, the high amount of work ($n = 10$) and the stress resulting from the obligation ($n = 10$) were mentioned as reasons against obligatory homework.

DA23_14 (translated): “The homework usually offered no added value for the exam, as they were often tasks that could not be implemented in the exam (e.g., creating a GeoGebra or Excel file, looking up the syllabus) [...] Since the obligatory homework is intended to prepare for the exam, it is wrong for ‘Didactics of Analysis’ and only creates additional time for the students, which is not used effectively for the exam.”

The students from this course also often referred to other courses ($n = 9$), in which they also had obligatory homework, and which were more important to them as they could gain more credits.

DA23_12 (translated): “We also have other courses with obligatory homework. These are more important because they earn more credits. It was far too much effort for so few credits.”

However, a few students in this course also mentioned positive arguments in favour of an obligation, such as a better working attitude ($n = 2$) or understanding ($n = 2$) and even a better preparation for the exam ($n = 4$).

Da23_38 (translated): “Obligatory homework motivates in a completely different way than no obligation. Obligatory homework means that the extensive knowledge is not learnt and forgotten in one go, but rather acquired over several weeks. In my opinion, anyone who has achieved 66% of the points in the homework independently should have no problems in the exam (with just a short repetition [of the content]).”

DISCUSSION

Although earlier studies (Neuhaus-Eckhardt & Siller, under review) showed that students in mathematics courses with obligatory homework rate this obligation rather positively, this is not the case for this mathematics education course. This may be due to the different content in mathematics and mathematics education courses and how the students view it because at least one student in this study also differentiated that homework in mathematics courses is needed to understand the content whereas it is not needed in mathematics education course.

The students in both courses with obligatory or voluntary homework assess an (potential) obligation in a similar way on the Likert-scale, but there are far more negative arguments regarding the course with obligatory homework. The students in this course confirm the high workload and stress that was already expected by their fellow students in the previous year and again mention other courses with obligatory homework, which seem to be more important as they can get more credits for these courses. In addition, many students in 2023 saw little relevance of the homework tasks especially for the exam, which may be caused by having to conduct the questionnaire after the exam. However, four students said that homework was helpful for the exam and some students in 2022 also mentioned to see no relevance for the exam. But in

contrast to 2023, at least three students in 2022 also stated to see relevance for their later profession as a teacher. However, as many students did not do the voluntary homework in 2022, this estimation could be based more on the content of the lecture than on the actual tasks.

Overall, the students' focus seems to be on achieving a good exam grade and on courses with as many credit points as possible, which is unsurprising but nevertheless worrying. In particular, at least some students seem to assume that homework is primarily intended to prepare them for the exam, so they request teaching to the test. The homework in this course was planned by the lecturer and assistants to deepen the students' understanding and address the integration of digital systems in mathematics lessons at school, which was communicated to the students. Typical exam tasks were also included into homework to prepare the students for their exam, but that was not the main reason for the often very time-consuming preparation of homework. However, if the students work on homework primarily with the aim of passing the exam, it can be assumed that they worked less on tasks which did not resemble exam tasks. In addition, only doing homework as preparation for the exam is an extrinsic motivation which is associated with lower learning outcomes (e.g., Ryan & Deci, 2000). It can therefore be assumed that these students did not build up a long-term understanding, which could be problematic for later courses or their later profession as a teacher.

In addition to their low utility values, the students also assigned high costs (and low benefit) to the homework, such as a high amount of work needed. Based on the situated expectancy-value theory (Eccles & Wigfield, 2020), they probably put little effort into completing the homework and inappropriate learning behaviours such as copying other solutions or only using ChatGPT were already mentioned by some students in their answers supporting Liebendörfer and Göller (2016) and Robinson and Kuin (1999).

When interpreting these results, it should be kept in mind that we only looked at a small sample from one type of course in mathematics education. Also, the homework was slightly varied in both courses to make copying more difficult. In addition, the survey in 2023 took place after the final exam, while in 2022 we surveyed the students before their final exam, although they had already seen exam like tasks. The different answers with references to the exam in 2023 may nevertheless be due to this reason. Furthermore, social desirability could have influenced the students' answers. Maybe some students thought that a negative assessment of the obligation would make homework voluntary again.

Further evaluations will analyze the extent to which homework has affected the students' performance in their exam. But the results of this study already revealed that many students did not see a relevance of the given homework, which could be considered while designing and implementing new homework. Maybe adding more tasks or using smaller tasks to trigger intrinsic motivation could be helpful to engage students in longer lasting learning behaviours. But some students may not prioritize

homework in a course at all, if they attend other courses with more credit points in parallel that require obligatory homework. Many students also mentioned that less stressful incentives, like bonus points for the final exam, could be a better motivator for doing homework than an obligation.

References

- Cooper, H. (1989). *Homework*. New York, NY: Longman.
- Eccles, J. S. & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation, *Contemporary Educational Psychology*, 61, 101859.
- Halcrow, C. & Dunnigan, G. (2012). Online homework in Calculus I: Friend or foe? *PRIMUS*, 22(8), 664–682.
- Koban, D., Fukuzawa M., Slocum, R., Fletcher, M. & Pleuss, J. (2020). Differential Effects of Incentivized Homework on Student Achievement in Undergraduate Mathematics. *PRIMUS*, 5(30), 501–519.
- Krapf, R., & Liebendörfer, M. (2021). Was bewirkt die Pflichtabgabe von Übungsaufgaben in der Hochschulmathematik? Ein empirischer Vergleich. (What is the effect of obligatory homework in university mathematics? An empirical comparison.) *mathematica didactica*, 44(2).
- Liebendörfer, M. & Göller, R. (2016). Abschreiben von Übungsaufgaben in traditionellen und innovativen Mathematikvorlesungen (Copying homework in traditional and innovative mathematics courses at university). *Mitteilungen der Deutschen Mathematiker-Vereinigung*, 24(4), 230–233.
- Mayring, P. (2022). *Qualitative Content Analysis. A Step-by-Step Guide*. London: Sage.
- Neuhaus-Eckhardt, S. & Siller, S. (2024). Homework at university - why do students (not) do assignments? *Proc. 47th Conf. of the Int. Group for the Psychology of Mathematics Education*. (Vol. 1, p. 190). PME.
- Neuhaus-Eckhardt, S. & Siller, H.-S. (under review). Homework at university – Relations to study success and students' characteristics.
- Pritchard, D. (2015). Lectures and transition: from bottles to bonfires? In A. C. Croft, M. J. Grove, J. Kyle, & D. A. Lawson (Eds.), *Transitions in Undergraduate Mathematics Education* (pp. 57–69). University of Birmingham.
- Robinson, V. M. J., Kuin, L. M. (1999). The explanation of practice: Why Chinese students copy assignments. *International Journal of Qualitative Studies in Education*, 12(2), 193–210.
- Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25, 54–67.
- Ufer, S. (2015). The role of study motives and learning activities for success in first semester mathematics studies. In K. Beswick, T. Muir, J. Wells, J. (Eds.), *Proc. 39th Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 4, pp. 265–272). PME.

EXPLORING THE ROLE OF INTERNALISED STANDARDS IN COMPARATIVE JUDGEMENT IN MATHEMATICS

Jennifer Palisse¹, Mark MacLean², Deborah King¹

University of Melbourne¹, University of British Columbia²

Comparative judgment has demonstrated positive effects on student learning, though the reasons why remain unclear. This study analyses think-aloud comments from students ($N = 24$) evaluating peer mathematics solutions involving rational inequalities, either comparatively or individually. Results suggest comparative judgment encourages students to focus on identifying commonalities and differences between solutions which reduces reliance on long-term memory retrieval. In contrast, sequential judgment requires recalling and applying internalised standards, making it more resource intensive. We argue that comparative judgement may be effective for learning because it reduces cognitive load.

LEARNING THROUGH COMPARATIVE JUDGEMENT

Comparative judgement is increasingly being used as an alternative assessment method that does not rely on traditional rubrics or marking guides (Jones & Karandeniz, 2014). Additionally, comparative judgement is being applied as a peer-review tool, where students evaluate, provide feedback, and rank pairs of their peers' work. This involves presenting pairs of work, side-by-side, and asking students to decide which of the two is 'better.' This is typically done without providing students with a rubric, and how students decide which solution is 'better' is up to them. Existing studies have found comparative judgement to be effective in improving performance outcomes in various educational contexts, including Design and Technology (Bartholomew, Strimel, et al., 2019), essay writing (Bouwer et al., 2018), and mathematics (Palisse et al., 2024).

One way to understand the effectiveness of comparative judgement is through the lens of analogical reasoning. Analogical reasoning involves identifying common patterns and relationships between problems, allowing learners to map similarities or differences from a known source problem to a new target problem (Gray & Holyoak, 2021). In some instances, a mental representation of these commonalities or differences may be stored in memory as an abstract schema, which can then be drawn upon and applied when solving similar problems. Comparative judgement operates in a similar way: by prompting learners to compare two worked solutions, it encourages them to identify and evaluate comparable similarities and differences. For example, a learner might compare two solutions and notice that one solution simplifies the fraction $\frac{8}{12}$ to $\frac{2}{3}$ by dividing both the numerator and denominator by 4, while another solution simplifies the same fraction by first dividing by 2 to get $\frac{4}{6}$ then by 2 again to get $\frac{2}{3}$. The

learner might then select the first solution as ‘better’ because it is more efficient. By noticing that each approach achieves the same result but with a different number of steps, the learner forms a mapping between the two solutions and identifies an *alignable difference*, where an alignable difference refers to a comparable element that exists between the two solutions but executed differently. We argue that the process of comparing the two solutions encourages learners to actively search for these commonalities and differences. It is these alignable differences that then forms the object to be evaluated. However, analogical reasoning rarely occurs spontaneously, and learners often fail to notice analogies on their own. Explicit prompts that highlight structural alignment between problems are necessary to aid learners in drawing inferences (Gentner et al., 2003). This is where comparative judgement’s explicit prompt to compare becomes important. By directly asking learners to make and justify a pairwise evaluation, it provides the kind of explicit scaffolding that encourages learners to actively search for patterns between solutions.

While the learning theory of analogical reasoning helps explain how comparative judgement might support learning, examining student perspectives offers additional empirical support for its effectiveness. Self-reported data indicates students enjoy taking part in comparative judgement (Bartholomew, Strimel, et al., 2019), often enjoying the process more than traditional face-to-face peer feedback (Bartholomew, Zhang, et al., 2019). Students appreciate the exposure to multiple pieces of work, which broaden their understanding of potential approaches and facilitate self-reflection (Strimel et al., 2020). Additionally, students report that comparative judgement helps them develop criteria for proficiency and become better at articulating what constitutes ‘high-quality’ work (Bartholomew, Zhang, et al., 2019). This process then promotes metacognitive skills and self-regulation, enabling students to internalise standards of quality and apply them to their own work (Bartholomew et al., 2018).

CURRENT STUDY

The current study extends previous findings (Palisse et al., 2024), which demonstrated that students who engaged in comparative judgement outperformed those who evaluated peer work sequentially. However, the mechanisms underlying this improvement remain unclear. While comparative judgement is often credited with improving performance outcomes as a result of explicitly comparing, Buckley et al. (2022) caution against prematurely attributing its benefits solely to this factor. They argue that positive outcomes may arise from a range of related influences, such as increased exposure to peers’ work, opportunities for critique, or the application of evaluative standards. Without a deeper understanding of the cognitive processes at play, it is difficult to pinpoint exactly why comparative judgement has been observed to be effective. This study investigates possible cognitive processes as a way of better understanding specific mechanisms that may make comparative judgement effective for learning.

METHOD

This paper forms part of a larger pretask-intervention-posttask study. We report here only on aspects of the research methods relevant to the current research question and analysis.

Participants

The study included 24 participants, with 15 being Year 10 (accelerated) and Year 11 students from a select entry secondary school in Victoria, Australia, all studying the same Year 11 mathematics subject. The remaining 9 were undergraduate students studying mathematics. Secondary students were new to rational inequalities, while the undergraduate students may have encountered rational inequalities sometime during their studies, but were out of practice.

Design

The think-aloud method (Ericsson & Simon, 1993) was used to capture students' thoughts while evaluating worked solutions. For the intervention, students were randomly assigned to one of two conditions: compare or sequential. Students in the compare condition were shown samples of worked solutions in pairs and asked to judge which of the two they felt was 'better'. Students in the sequential condition were shown the same set of worked solutions one-at-a-time and asked to assign each a score out of 5. Each condition had 12 students, with an equal distribution of secondary and undergraduate students.

Intervention

Students evaluated a set of eight worked solutions to the following problem: Find the set of real numbers, x , such that $\frac{x+1}{x-7} > 3$, where $x \in \mathbb{R} \setminus \{7\}$. Worked solutions included (1) correct and incorrect answers; (2) a range of methods and solution approaches from algebraic to graphical; (3) both high- and low-quality work; and (4) neat and messy solutions.

Procedure

All data collection occurred in a single problem-based interview lasting between 45 and 60 minutes. Students evaluated the worked solutions, either in pairs or one-at-a-time. Students who compared solutions were asked to form a judgement for each pair based on three prompts: (1) quality of mathematical understanding; (2) quality of communication; and (3) overall quality. Students in the sequential group rated solutions on the same criteria and assigned scores out of five. All students were shown the same set of eight solutions and in the same order for their respective experimental conditions. Students were asked to think-out-loud while evaluating and were informed that there were no correct or incorrect judgements, that the way in which they evaluated each solution was up to them. Students were not provided with marking schemes or correct answers. After each round of evaluation, the worked solution(s) were removed so that students could no longer see them. The think-aloud method was used as the

primary tool to access students' conscious thoughts during intervention. Think-aloud data were audio recorded and transcribed.

RESULTS

Results indicate that students engaged in predominantly two approaches to evaluative judgement, which we will refer to as *comparative judgement* and *internalised criteria-based judgement*. We provide illustrative examples for both approaches, chosen for their clarity and representation of common student experiences. An analysis of the prevalence of each evaluative approach is then presented, noting differences between the two experimental conditions.

Internal criteria-based judgement

First, two students who evaluated solutions individually are introduced along with examples of these students applying internalised criteria-based judgements. The first is Owen, who evaluates a solution that uses a method he feels lacked rigor.

Owen: Well, the answer seems to make sense. I was also taught this method in high school. But I don't like it. In order for this method to work, you have to prove that the function satisfies certain conditions, I think. But that would be more complicated. But I guess most of the markers would give credit for that. Well, it's correct. Except for the fact that they might not understand the method that they were using.

Owen activates prior knowledge, recalling learning this method in high school. He shows awareness of different evaluative standards, acknowledging that what he finds acceptable might differ from what subject assessors might find acceptable, indicating that Owen differentiates between commonly accepted standards and his own preferences. He applies his personal standard of expecting mathematical rigour ("you have to prove that the function satisfies certain conditions") and uses this to assess the worked solution. He criticises the solution because it does not meet the expectations of his standard ("But I don't like it."). The process of Owen's evaluative judgement is as follows: the worked solution triggers relevant prior knowledge and experiences, drawn from long-term memory, then applied to the worked solution to make a judgement.

In the second example, Eugene evaluates a solution to a rational inequality problem using a number line. However, he would prefer the number line to be represented differently.

Eugene: This looks fine for now. With the number line... actually, I would prefer to shade the overlapping domains of the number lines. That's my preference to show the examiner 'Oh, I want this domain for my x .' But that's just me.

Eugene acknowledges that his judgement is based on personal preferences rather than a necessary mathematical standard ("But that's just me."). This further highlights how the evaluative process is guided by internalised standards that are unique to one's own experiences. Specifically, his preference for number lines to include shading to represent overlapping domains reflects his own idealised preference for how number lines should be represented. During the evaluative judgement process, Eugene activates

this stored mental representation and holds it in working memory. This then forms a benchmark against which to assess the solution in front of him. Eugene forms his judgement by comparing the actual solution against his idealised version, where the inclusion or lack of shading becomes an alignable difference between the two, and concludes that his idealised version executes this feature more effectively.

Comparative judgement

Next, we turn to examples of comparative judgement, where two students who evaluated solutions in pairs formed evaluations by identifying differences between solutions.

The first example is Mabel, who compares two solutions with contrasting layouts. In one solution, a graph is positioned in the centre of the page, while the accompanying algebraic working is located in the corner, creating a sense of disconnection between the two. In contrast, the second solution presents all lines of working sequentially in a single cohesive block. It is this difference in layout that catches Mabel's attention.

Mabel: This one (Left) is a bit more messy. A bit more all over the place. This one (Right) you can follow it from top to bottom.

Mabel's evaluative judgement begins with her recognition of a difference in layout and organisation. This alignable difference – the structure of the working – becomes the focus of her comparison. To form her judgement, Mabel retrieves an internalised standard, valuing clarity and ease of understanding. She then applies this standard to the noticed difference in layout, later choosing the second solution as 'better' as it aligns more closely with her expectation of easy-to-follow working ("you can follow it from top to bottom."). Thus, Mabel's evaluative judgement process involves noticing an alignable difference and then drawing upon her internalised preference for clarity to decide which solution meets this standard more effectively.

The second example is Norman, who compares a solution using a sign diagram to another that uses an algebraic method but does not provide a final answer. Despite struggling to fully understand the mathematics in both solutions, Norman identifies a key difference between them.

Norman: I don't know what a sign graph is. Yeah, I still don't really understand this graph. But the sign graph one has come to an answer, so I think this one is better. Although I don't understand it, it has all the steps of what they're doing, and they write what they're doing. But the other one confused me. I'm not sure what the answer, or conclusion, is. Because I don't think they've finished it. They've just left it.

Norman's example demonstrates how students can engage in comparative judgement without requiring a full understanding of the mathematics. This aligns with findings from Jones and Alcock (2013), who showed that social science postgraduate students, expected to lack the mathematical knowledge to understand undergraduate calculus solutions, were still able to form comparative judgements that correlated with those of mathematics lecturers.

Norman's evaluative judgement process involves identifying an alignable difference between the two solutions – one includes a final answer, while the other does not. This is the element he will then judge. To evaluate this element, Norman draws upon an internalised standard that values completeness (“it has all the steps”). Norman then selects the solution that better executes ‘completeness’ as the ‘better’ solution.

Distribution of evaluative judgement approaches

Having examined how individual students approached evaluation through the case studies, we now extend our analysis to all 24 participants to understand the broader distribution of criteria-based and comparative judgements. Table 1 shows the number of criteria-based and comparative judgements made by students from each group.

Group	N	Internal judgement	Comparative judgement
Sequential	12	212	34
Compare	12	74	129

Table 1: Number of internalised criteria-based judgements and comparative judgements generated by students

While both groups used both types of judgement, their distribution differed significantly, $\chi^2(1) = 116.79, p < 0.001$. The sequential group engaged in approximately six times more criteria-based judgements than comparative judgements. Conversely, the compare group favoured comparative judgements over criteria-based judgements. However, the total number of judgments made did not differ significantly between groups, $t(14) = 0.82, p = 0.425$. These findings demonstrate that the way solutions are presented – either comparatively or individually – does not influence the number of judgements students make, but rather how students approach evaluative judgement.

DISCUSSION

This study examined why comparative judgement might be effective for learning, identifying two distinct approaches to evaluative judgement: comparative judgement and internalised criteria-based judgement. In comparative judgement, students identify alignable differences between two solutions, then apply an internalised standard to evaluate which solution better addresses the identified difference. In contrast, internalised criteria-based judgement involves noticing elements of a solution that trigger prior knowledge or experiences, from which evaluative standards are retrieved from long-term memory and applied to form a judgement. The presentation of solutions strongly influenced the approach students used. Paired solutions encouraged comparative judgement, while individual solutions encouraged internalised criteria-based judgement.

These distinct approaches have implications for the effectiveness of learning when using comparative judgement. To better understand how these two approaches differ in their cognitive demands, we can contrast their key processes.

Comparison: During comparative judgement, both solutions are available to the learner, physically in front of them, so that neither solution needs to be recalled and held in working memory. During internalised criteria-based judgement, the learner assesses the solution in front of them against a benchmark drawn from long-term memory. This benchmark is recalled and held in working memory during the evaluation, where the learner identifies similarities or differences between the benchmark and the actual solution.

Identifying elements to evaluate: Comparative judgement involves identifying patterns or differences between solutions. These patterns then become the focus of evaluation. Criteria-based judgement instead requires searching for elements that trigger associations with internalised standards.

Use of long-term memory: Both approaches draw on long-term memory, but at different stages. During comparative judgement, after identifying elements of interest, the next step is to draw upon internalised criteria to form a judgement in order to decide which elements are executed better in each solution. During criteria-based judgement, standards stored in memory guide evaluations from the start.

These differences between internalised criteria-based judgement and comparative judgement highlight why presenting solutions as pairs may be more effective than presenting them individually, drawing parallels to analogical reasoning. In analogical reasoning, when both the source and target problems are provided, learners can focus on mapping one to the other. However, if only the target problem is given, learners face the more challenging task of first searching long-term memory for a relevant source, retrieving it, and holding it in working memory while simultaneously mapping it to the target problem (Thagard et al., 1990). A similar distinction may be occurring in evaluative judgement. Comparative judgement, where two solutions are physically present, allows learners to focus on mapping corresponding elements between solutions, akin to having both the source and target problems available. In contrast, internalised criteria-based judgement mirrors the more demanding case where only the target problem is provided. Learners must retrieve relevant standards from memory, maintain them in working memory, and compare them against the solution. This suggests that presenting solutions individually imposes greater cognitive demands.

Future research could investigate whether the hypothesis that internalised criteria-based judgement is indeed more resource-intensive by measuring cognitive load. Additionally, the case of Norman showed that comparative judgement is still an achievable process for those with low prior knowledge. Further research could look at individual learner differences and the relationship between prior knowledge and the effectiveness of comparative judgement versus internalised criteria-based judgement. Studies could explore factors such as working memory capacity, cognitive flexibility, or prior domain knowledge to see how these interact with the two evaluative judgement approaches.

REFERENCES

- Bartholomew, S. R., Strimel, G. J., & Yoshikawa, E. (2019). Using adaptive comparative judgment for student formative feedback and learning during a middle school design project. *International Journal of Technology and Design Education*, 29, 363–385. <https://doi.org/10.1007/s10798-018-9442-7>
- Bartholomew, S. R., Yoshikawa, E., & Connolly, P. (2018). Exploring the Potential for Identifying Student Competencies in Design Education through Adaptive Comparative Judgment. In N. Seery, J. Buckley, D. Canty, & J. Phelan (Eds.), *36th International Pupils' Attitudes Towards Technology Conference* (pp. 187–194).
- Bartholomew, S. R., Zhang, L., Bravo, E. G., & Strimel, G. J. (2019). A Tool for Formative Assessment and Learning in a Graphics Design Course: Adaptive Comparative Judgement. *Design Journal*, 22(1), 73–95. <https://doi.org/10.1080/14606925.2018.1560876>
- Bouwer, R., Lesterhuis, M., Bonne, P., & De Maeyer, S. (2018). Applying Criteria to Examples or Learning by Comparison: Effects on Students' Evaluative Judgment and Performance in Writing. *Frontiers in Education*, 3, 86. <https://doi.org/10.3389/educ.2018.00086>
- Buckley, J., Seery, N., & Kimbell, R. (2022). A Review of the Valid Methodological Use of Adaptive Comparative Judgment in Technology Education Research. *Frontiers in Education*, 7. <https://doi.org/10.3389/educ.2022.787926>
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol Analysis: Verbal reports as data (revised edition)*. Bradford Books/MIT Press.
- Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology*, 95(2), 393–408. <https://doi.org/10.1037/0022-0663.95.2.393>
- Gray, M. E., & Holyoak, K. J. (2021). Teaching by Analogy: From Theory to Practice. *Mind, Brain, and Education*, 15(3), 250–263. <https://doi.org/10.1111/mbe.12288>
- Jones, I., & Alcock, L. (2013). Peer assessment without assessment criteria. *Studies in Higher Education*, 39(10), 1774–1787. <https://doi.org/10.1080/03075079.2013.821974>
- Jones, I., & Karandeniz, I. (2014). An Alternative Approach To Assessing Feasibility. *Proceedings of the 2016 40th Conference of the International Group for the Psychology of Mathematics Education*, 126–136.
- Palisse, J., King, D. M., & MacLean, M. (2024). Comparative judgement and its impact on the quality of students' written work in mathematics. *Fifth Conference of the International Network for Didactic Research in University Mathematics*, 788–799.
- Strimel, G. J., Bartholomew, S. R., Purzer, S., Zhang, L., & Ruesch, E. Y. (2020). Informing engineering design through adaptive comparative judgment. *European Journal of Engineering Education*, 1–20. <https://doi.org/10.1080/03043797.2020.1718614>
- Thagard, P., Holyoak, K. J., Nelson, G., & Gochfeld, D. (1990). Analog retrieval by constraint satisfaction. *Artificial Intelligence*, 46(3), 259–310. [https://doi.org/10.1016/0004-3702\(90\)90018-U](https://doi.org/10.1016/0004-3702(90)90018-U)

HOW DO CULTURAL NORMS INFLUENCE TEACHER NOTICING?

Josephine Friederike Paul¹, Anika Dreher², Ting-Ying Wang³, Feng-Jui Hsieh³ & Anke Lindmeier¹

¹FAU Erlangen-Nürnberg, Germany; ²PH Freiburg, Germany; ³NTNU, Taiwan

Teacher noticing is influenced by teacher characteristics and other factors like instructional norms, which can differ between cultural contexts. Moreover, it is questionable whether norms are shared between different expertise groups. Thus, in the TaiGer Noticing project, we used text vignettes to examine 36 Taiwanese and German mathematics education professors' perspectives on instructional situations and 221 Taiwanese and German teachers' noticing. The contribution focuses on one vignette, and the findings indicate an interculturally valid norm regarding the use of a task at the professor level. However, while a majority of the German teachers referred to this norm in their noticing, the Taiwanese teachers focused more on other aspects, for example, explanations. Possible reasons for these findings are discussed.

THEORETICAL BACKGROUND

Mathematics teachers' cognitive characteristics are relevant factors for high-quality mathematics instruction (e.g., Hill et al., 2008). Besides knowledge itself, its use, for example, in noticing processes, is central. Noticing is often considered the interplay of “attending to particular events in an instructional setting and making sense of those events” (Sherin et al., 2011, p. 9) based on beliefs and knowledge. It is usually investigated by vignettes representing instructional situations in video or text (Weyers et al., 2023). Teachers are asked to analyze (specific aspects of) the situations, and their answers are used to infer underlying noticing processes (Friesen & Dreher, 2024).

Besides knowledge and beliefs, instructional norms affect teachers' evaluations of instructional situations (e.g., Herbst & Chazan, 2011). Norms are shared convictions negotiated in a social group, with instructional norms referring to the social context of teaching. Instructional norms create behavior expectations (Herbst & Chazan, 2011), for example, how a teacher should react to a student's error. If one holds a norm and observes a situation where these expectations are not met, this will catch the observer's attention and result in a (critical) reaction. Thus, reactions are indicative of certain norms (breach of a norm approach, e.g., Herbst & Chazan, 2011).

As norms are negotiated in social groups, they may vary between such different groups. The norms and associated beliefs may lead to different noticing. Thus, noticing could differ between expertise groups, such as mathematics teachers as practitioners and mathematics education professors as researchers. Although professors may transmit some of their norms in their role as teacher educators (Paul et al., 2024), teachers may

not adopt all of them. This is often referred to as the theory-practice gap (e.g., Lindmeier et al., 2018). Moreover, negotiation processes, e.g., among teachers' colleagues or in class, and individual beliefs based on experiences can counteract previously internalized norms. However, teachers and teacher educators may share particular perspectives when knowledge translations, characterized by negotiation and adaption processes, succeed (e.g., Lindmeier et al., 2018; McNeal & Simon, 2000).

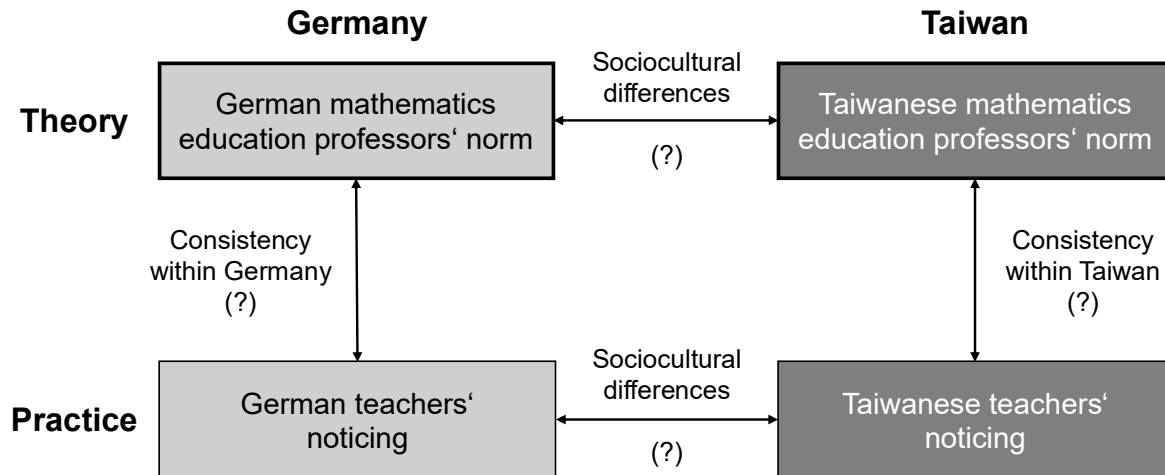


Figure 1. TaiGer Noticing framework for investigating the consistencies and differences of instructional norms between different cultures and expertise groups

Besides expertise groups, norms can vary between groups embedded in different sociocultural contexts, like Taiwanese and German professors. For example, varying cultural traditions, such as more product-oriented instruction in East Asian and more process-oriented instruction in Western countries, were described, which are related to instructional norms (Leung, 2001). Previous findings in the TaiGer Noticing project could evidence that Taiwanese and German mathematics education professors adhere to different norms when dealing with a student's mistake (Dreher et al., 2021): While the majority of Taiwanese professors addressed finding only one solution when solving a quadratic equation as incorrect, German professors saw the partial solution as potential for further learning (Dreher et al., 2021). However, regarding other aspects, Taiwanese and German mathematics education professors share certain norms (e.g., Paul et al., 2024). One open question is how these commonalities and differences are reflected in the noticing of teachers when considering different cultural contexts.

RESEARCH INTEREST

To date, research on culture-dependent instructional norms and their influence on teacher noticing is still scarce. While divergent instructional traditions (e.g., in Taiwan and Germany) and the theory-practice gap (between professors and teachers) suggest that perspectives vary between the groups in focus, knowledge translations and international exchange suggest that teacher noticing reflects professors' norms and that groups from different countries would share perspectives (see the illustration in Figure 1). To clarify the different relations, the TaiGer Noticing, the TaiGer Noticing project investigates instructional norms through Taiwanese and German professors' noticing

and examines whether and how teachers' noticing reflects the identified norms. This report focuses on the research questions: RQ1) Does the Taiwanese and German mathematics education professors' noticing indicate an interculturally shared norm when evaluating the use of a task aimed at promoting flexible solving? RQ2) If so, is the norm equally reflected in Taiwanese and German mathematics teachers' noticing?

METHODS

The analysis is part of the Taiwanese-German project "TaiGer Noticing". It aims at the investigation of how cultural norms influence teacher noticing. In the project's first phase, Taiwanese and German mathematics education professors' evaluations of instructional situations were investigated to explicate a frame of reference for norms that could influence teacher noticing (e.g., Paul et al., 2024 for the detailed findings of the first project phase). In the second phase, we used the standard defined by the professors' answers to examine teacher noticing regarding the same situations.

Sample


The data from the first phase of the TaiGer Noticing project used for this analysis included written evaluations of 19 Taiwanese and 17 German mathematics education professors. All of them were active in secondary mathematics teacher education and mathematics education research. In the project's second phase, 124 German and 115 Taiwanese mathematics teachers participated. However, the data of 7 German and 11 Taiwanese teachers were missing, incomplete, or off-topic and hence excluded. Thus, the dataset comprises written evaluations of 117 German (58 female, 53 male, 6 missings) and 104 Taiwanese mathematics teachers (53 female, 48 male, 3 missings). German teachers had, on average, 12.2 years of teaching experience (after initial in-service training, so-called *Referendariat*), and Taiwanese teachers 17.1 years.

Vignette

The participants evaluated the use of a task as represented in a text vignette (Figure 2). This text vignette was originally authored in Germany but developed in a bicultural concurrent process (Dreher et al., 2021). It consists of 1) a short description of the situation and 2) a picture of a task, which, from the authors' perspective, has the potential to foster flexible solving. Moreover, it includes a 3) fictive teacher-student dialogue mentioning different solution strategies. The study participants were asked to "Please evaluate the teacher's use of the task in this situation and give reasons for your answer" (Paul et al., 2024). A breach of the norm anticipated by the authoring team was: although three different solution strategies are mentioned, the teacher addresses only one in detail and notes it on the blackboard.

In the first phase of the TaiGer Noticing project, the vignette was presented to German and Taiwanese mathematics education professors (RQ1). In the second phase, the vignette was presented to the teachers (RQ2). The participants read and evaluated the vignette in their native language. Subsequently, the written evaluations were translated into English, the research team's common language.

The teacher T asked the students to do the following task from a textbook as homework to promote their ability to *flexibly solve real-life applications*:

From a children's fever medicine prescription:	
Dosage of this drug depends on weight! For every 5 kg of body weight, take 15ml of the medicine.	
Paul weights 17.5 kg. What dosage does he need? This problem can be solved with different strategies. Can you find at least two?	

T: Let's talk about the homework task now. How did you do it?

S1: I just read the problem and looked at the numbers. You need 3 and a half times the 15 ml. This was easy.

T: Good job, S1. How about someone else? What strategies did you use?

S2: I made a table for a few weights. The solution is 52.5 ml.

T: Well done! So we found already two ways to solve the problem: Looking at the numbers or using a table. Does anyone have any other ideas?

[Many students raise their hands.]

S3: I figured out the equation for the proportional relationship and used it for the calculation.

T: Ok great! So you used the function that relates the kg to the ml. I would like to put a solution on the blackboard so that all of you can compare it with your own solutions. S3, will you please carefully describe what you did?

S3: Yes. To determine the constant of the proportion you divide 15 ml by 5 kg. So you get 3 ml per kg. And then you can set up the equation $y = 3x$. So you have $y = 3 \cdot 17.5 = 52.5$. [T writes the solution on the blackboard.]

T: So 52.5 ml is indeed the appropriate dosage for Paul. Well done, everyone.

Figure 2. The vignette consisting of a short description of the situation, a picture of a word problem, and teacher-student dialogue (Paul et al., 2024)

Coding

To answer RQ1, professors' evaluations were coded in two steps (qualitative content analysis; Mayring, 2014). In the first step, evaluations were categorized as including or not including critical aspects regarding the use of the task (*critique* vs. *no critique*). In the second step, the evaluations that included *critique* were further coded to capture the reasons for the critique. A deductive category, "critique regarding the anticipated breach of a norm" (*N*), was implemented. Other reasons for critique were identified inductively. Thus, two further categories were developed: critique regarding the explanations (*E*) and other reasons only mentioned by individual persons (*O*). *E* was coded when participants "considered the explanation of the solution paths insufficient" (Paul et al., 2024). While *N* and *E* could be coded together, *O* was coded for answers that did not fit the before-mentioned categories, so only when neither *N* nor *E* was assigned. Details on coding the professors' answers can be found in Paul et al. (2024).

In the project's second phase, teachers' evaluations were coded using the manual developed in the first phase. Each research team member individually coded evaluations of 20 German and 15 Taiwanese teachers, followed by an independent consensus coding within the German resp. Taiwanese teams. The two codings of the 35 evaluations were then discussed between the teams to establish a common

understanding. Subsequently, the entire dataset was coded once in Germany and once in Taiwan. Interrater reliability was assessed using a binary string (0 = code not assigned, 1 = assigned) for the codes *no critique*, *N*, *E*, and *O*. The code *critique* was not included as it implicitly is included in the other codes. Cohen's Kappa calculated for approx. 80% of all evaluations was .68. The entire research team then conducted a consensus coding.

Procedures of analysis

Based on the final dataset, the proportions of professors and teachers mentioning certain critical aspects were calculated and compared between groups, as visualized in Figure 1. Per country, when more than 50% (the majority) of the professors, who were experts in their field, criticized the breach of the anticipated norm, it was considered an actual norm in their context (Dreher et al., 2021). Moreover, χ^2 -tests were conducted to test the significance of the differences in code distributions between the subsamples.

FINDINGS

Table 1 shows the proportions of mathematics education professors and teachers giving specific reasons for critique. While the proportions of German and Taiwanese professors expressing *critique* were similar (around 70%), the German teachers were significantly more critical than Taiwanese teachers ($\chi^2 = 17.32, p < .001$).

		Professors		Teachers	
		Germany (N=17)	Taiwan (N=19)	Germany (N=117)	Taiwan (N=104)
<i>Critique</i>		12 (70.6%)	13 (68.4%)	89 (76.1%)	51 (49.0%)
as anticipated regarding different strategies	<i>N</i>	10 (58.8%)	11 (57.9%)	75 (64.1%)	28 (26.9%)
regarding explanations	<i>E</i>	6 (35.5%)	3 (15.8%)	18 (15.4%)	22 (21.2%)
other	<i>O</i>	0 (0%)	1 (5.3%)	6 (5.1%)	9 (8.6%)

Table 1. Proportions of answers issuing *critique* and giving different reasons for critique (*N*, *E*, *O*) for all subsamples. Bold: Majorities indicative of the norm.

This is even stronger when only considering the critique regarding the anticipated breach of the norm (*N*): Most German and Taiwanese professors mentioned a breach of the anticipated norm, making it an interculturally valid norm (bold in Table 1, see Paul et al., 2024). However, the German teachers' noticing significantly more often reflected this norm compared to their Taiwanese counterparts ($\chi^2 = 30.58, p < .001$). Taiwanese teachers who did not criticize anything often mentioned positive aspects regarding student-centeredness or gave descriptive answers. Although teachers' noticing reflected the norm determined by the professors, the answers showed certain differences: While professors explained quite comprehensively how solution paths should have been compared or related, some teachers just shortly criticized the notation of only one instead of all solution strategies, as in the following example:

German teacher: “Due to various solution possibilities, it makes sense to use this task. However, for the promotion of flexible solving, **all solution paths** should also be written down.” [emphasis added for this report]

Taiwanese teachers, whose evaluations included a reference to the professors’ norm (coded with *N*), often provided more detailed explanations and interpretations than their German counterparts. They elaborated on how the strategies were related or should have been related. For example, this is evident in the following evaluation of a Taiwanese teacher who addressed the comparison of “differences and similarities between [...] approaches”. However, the German *and* Taiwanese teachers’ answers used as illustrations in this report highlighted the insufficiency of addressing only one solution path, so both adhered to the norm to *address several paths when aiming at flexible solving* (validated on the professor level).

Taiwanese teacher: “It would be better to have students go up to the board and write things down so that they **explain things more clearly** as they are presenting instead of just touching upon the topic. Having students present their solution allows other students to find out what other solutions there are and to think further about whether they are all feasible. In addition, it would be better to **compare the differences and similarities between those different approaches**, which use different concepts, so that students have a better understanding of those approaches.” [emphasis added for this report]

In the Taiwanese teacher’s answer, it is also observable that the participant puts special emphasis on students’ understanding, which was not grasped by any of the codes developed in the first project phase. Moreover, code *E* was assigned to this answer as the teacher requested the students to “explain things more clearly.” Similar critiques regarding the explanations were found in 15 – 36 % of the evaluations in the four subsamples. Participants criticized that the algebraic characteristics of the solution paths should have been explained more thoroughly or that the teacher should have questioned the solution paths more. While more German than Taiwanese professors addressed the aspect of insufficient explanations (*E*), a reverse pattern was observed for the teachers. Moreover, there was a significant discrepancy between German professors and teachers, with professors mentioning this point of critique more than twice as likely ($\chi^2 = 4.00, p < .05$). Other reasons for critique (*F*) were expressed by approx. 9% of the Taiwanese and 5% of the German teachers.

DISCUSSION

The study aimed to investigate whether instructional norms influence teacher noticing. Before discussing the findings, we remind that instructional norms are social norms and reflect what is expected within a group rather than an objectively best approach.

The analyses on the professor level showed that more than 50% of German and Taiwanese mathematics education professors criticized the breach of the anticipated norm of *equally addressing different solution paths when aiming at flexible solving*. Thus, this norm can be considered an interculturally valid one (Paul et al., 2024).

Comparing Taiwanese and German teachers' noticing on this basis, more German teachers referenced the norm violation. Yet, more Taiwanese teachers who criticized the breach of the norm showed more in-depth noticing than their German counterparts. Thus, although the noticing of the German teacher sample occurred to be more in line with the norm determined by the professors' perspectives (more coherence), several Taiwanese teachers who criticized the breach of the norm provided more elaborate interpretations as part of their noticing. This emphasizes the importance of how we operationalize noticing and which indicators we consider (e.g., attending to, interpretation) when assessing noticing. Moreover, we found that Taiwanese teachers focused more on the details of the explanation (E) and positive aspects, such as student-centeredness. This observation aligns with prior descriptions that characterize East Asian mathematics instruction as effective instruction when emphasizing reflection, understanding, and product orientation (e.g., Leung, 2001; Pratt et al., 1999).

Comparing professors and teachers, we found that the teachers' evaluations were more variable than the professors'. This could be due to teachers' more diverse beliefs. Also, professors were more critical than teachers, which may be attributed to their role in maintaining high standards within their field of expertise. Yet, while on the professor level, no significant differences between countries showed regarding the critique, German teachers were more likely to criticize aspects of task use. The latter might reflect societal differences, such as German culture's typically more critical mindset.

Before concluding our findings, we want to address some limitations briefly. First, the findings reported are based on one vignette as an example to illustrate the differences and consistencies between German and Taiwanese professors and teachers, limiting its scope to the presented vignette. Second, only the codes from the first phase were used to code the teachers' answers. More in-depth analyses regarding the answers coded with *no critique* or *O* need to be conducted to trace more subtle differences in noticing.

Nevertheless, the analyses showed two things. First, when considering the norm interculturally shared on the professor level as a standard, our findings suggest that the German teachers noticed better in the sense of being more aligned with the standard. However, several German teachers who referred to this norm gave less detailed answers than their counterparts. Second, our findings illustrate that while mathematics teachers' noticing aligns with certain expert (professor) norms in one country, this does not have to apply in another cultural context. One reason for a theory-practice gap might be differences in knowledge or beliefs between different expertise groups. Nevertheless, future research is needed to investigate this in detail.

To summarize, the findings emphasize that even when using cross-culturally shared norms at the professor level as a standard, substantial differences in teachers' noticing may occur. This underlines that researchers must carefully decide on standards used in cross-cultural noticing research. Thorough pilot phases and systematic investigations with expert and teacher samples may be an advised means to inform these decisions.

Acknowledgments

This study is part of the TaiGer Noticing project, funded by MOST (grant nr. 106-2511-S-003-027-MY3) and the DFG – (grant nr. DR 1098/1-1 and LI 2616/2-1).

References

- Dreher, A., Lindmeier, A., Wang, T.-Y., Feltes, P., & Hsieh, F.-J. (2021). Do cultural norms influence how teacher noticing is studied in different socio-cultural contexts? A focus on expert norms of responding to students' mathematical thinking. *ZDM*, 53(1), 165–179.
- Friesen, M., & Dreher, A. (2024). Teacher noticing regarding aspects of instructional quality in the mathematics classroom. In R. Stahnke & A. Gegenfurtner, *Teacher Professional Vision: Empirical Perspectives* (Vol. 1, p. 101–115). Routledge.
- Herbst, P. G., & Chazan, D. (2011). Research on Practical Rationality: Studying the justification of actions in mathematics teaching. *The Mathematics Enthusiast*, 8(3), 405–462.
- Hill, H. C., Blunk, M. L., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L., & Ball, D. L. (2008). Mathematical Knowledge for Teaching and the Mathematical Quality of Instruction: An Exploratory Study. *Cognition and Instruction*, 26(4), 430–511.
- Leung, F. K. S. (2001). In Search of an East Asian Identity in Mathematics Education. *Educational Studies in Mathematics*, 47, 35–51.
- Lindmeier, A., Riecke-Baulecke, T., & Barzel, B. (2018). Berufsbegleitende Lehrerbildung als Profession verstehen – Konzeption eines Weiterbildungsmasterstudiengangs für Fort- und Auszubildende von Mathematiklehrpersonen. In R. Biehler, T. Lange, T. Leuders, B. Rösken-Winter, P. Scherer, & C. Selzer (eds.), *Mathematikfortbildungen professionalisieren* (p. 435–452). Springer Fachmedien Wiesbaden.
- Mayring, P. (2014). *Qualitative Content Analysis: Theoretical Foundation, Basic Procedures and Software Solution*. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-395173>
- McNeal, B., & Simon, M. A. (2000). Mathematics Culture Clash: Negotiating New Classroom Norms with Prospective Teachers. *The Journal of Mathematical Behavior*, 18(4), 475–509.
- Paul, J. F., Dreher, A., Wang, T.-Y., Hsieh, F.-J., & Lindmeier, A. (2024). Culture-Specific Norms Regarding High-Quality Use of Task Potential for Mathematical Learning—Contrasting Researchers' Perspectives from Germany and Taiwan. *Journal Für Mathematik-Didaktik*, 45(2), 14.
- Pratt, D. D., Kelly, M., & Wong, W. S. S. (1999). Chinese Conceptions of 'Effective Teaching' in Hong Kong: Towards Culturally Sensitive Evaluation of Teaching. *International Journal of Lifelong Education*, 18(4), 241–258.
- Sherin, M. G., Jacobs, V. R., & Philipp, R. A. (2011). Situating the Study of Teacher Noticing. In M. G. Sherin, V. R. Jacobs, & R. A. Philipp, *Mathematics Teacher Noticing*. Routledge.
- Weyers, J., König, J., Santagata, R., Scheiner, T., & Kaiser, G. (2023). Measuring teacher noticing: A scoping review of standardized instruments. *Teaching and Teacher Education*, 122, 103970.

MATHEMATICS SPECIALIZED KNOWLEDGE OF SPECIAL EDUCATION TEACHERS TO TEACH NUMBERS AND OPERATIONS TO STUDENTS WITH VISUAL IMPAIRMENTS

Juan Luis Piñeiro¹, Carola Manolino^{2,3}, Eder Pinto⁴ and Catalina Román⁵

¹Universidad Metropolitana de Ciencias de la Educación; ²Dipartimento di Scienze Umane e Sociali, Università della Valle d'Aosta; ³Laboratorio "S. Polin", Dipartimento di Matematica "G. Peano", Università di Torino; ⁴Instituto de Ciencias de la Educación, Universidad de O'Higgins; ⁵Facultad de Ciencias Sociales, Universidad de Chile

This paper describes the mathematics specialized knowledge mobilized by three special education teachers when teaching natural numbers and arithmetic operations in a school for students with visual impairments. The results show that instruction is delivered through a guided, question-based, step-by-step approach focused on procedures. The study concludes with a call to mathematics and special education teacher educators to address the mathematical preparation of this group of teachers from a cultural and formative perspective. This involves adopting a reflective and conceptual approach to mathematics rather than a purely instrumental one.

Mathematics instruction in special education has often relied on strategies that tend to discourage the development of students' mathematical thinking (e.g., Tan et al., 2020). Explicit teaching, for instance, has been described as one of the strategies with immediate results suited to students with mathematical learning difficulties (e.g., Gersten et al., 2009). However, this strategy—characterized by guided, step-by-step methods—often prevents students from engaging in critical thinking and conceptual understanding. While explicit instruction can provide structure, it limits opportunities for students to delve into mathematical ideas or engage in problem-solving processes. Instead, with this method, students are confined to only repeating recipes (Lambert, 2018).

A key factor in addressing this issue lies in the knowledge of the teacher who teaches mathematics. This construct is understood as the professional knowledge required to teach and is critical in fostering the learning of school mathematics (e.g., Blömeke et al., 2022). However, special education teachers—who usually teach to visual impairments students—have been excluded from research exploring the specific knowledge required to teach mathematics. In this context, literature presents models of knowledge that fail to effectively integrate the visual impairments expertise with disciplinary knowledge, such as mathematics.

This paper focuses on special education teachers working with students with visual impairments. Prior research suggests that the knowledge of this group of teachers includes teaching specialized calculation strategies (Klingenberg et al., 2019), i.e.,

mathematical signography, the Japanese Soroban abacus, or the Braille typewriter (Brawand & Johnson, 2016). However, limited understanding exists regarding special education teachers' knowledge to teach mathematics (Allsopp & Haley, 2015), particularly to students with visual impairments. This context raises the following question: What specialized knowledge do special education teachers mobilize when teaching natural number and arithmetic operations to students with visual impairments?

THEORETICAL PERSPECTIVE

The literature on the knowledge of special education teachers has seen sustained development in recent years, though it remains insufficient to fully understand the mathematical knowledge required in this context (Allsopp & Haley, 2015; van Garderen et al., 2013). Advances in this field have mainly focused on pedagogical content knowledge (e.g., Firestone et al., 2021), yet they remain fragmented and inconsistent in their conceptualization and application to disciplinary content. This study adopts a theoretical perspective rooted in mathematics education research to address the research question. Specifically, it draws on the Mathematics Teacher Specialized Knowledge (MTSK) model developed by Carrillo et al. (2018). This theoretical framework enables the description of teacher's knowledge and the examination of skills mobilized in mathematics classrooms with students with visual impairments. As Carrillo et al. (2018) point out, "this model features a reconfiguration of mathematical knowledge, a reinterpretation of pedagogical content knowledge and a new way of conceptualizing the notion of specialization" (p. 240), considering two domains of knowledge: Mathematical Knowledge (MK) and Pedagogical Content Knowledge (PCK). Each domain is further divided into subdomains. For MK, these include Knowledge of Topics (KoT), Knowledge of the Structure of Mathematics (KSM), and Knowledge of Practices in Mathematics (KPM). For PCK, the subdomains are Knowledge of Mathematics Teaching (KMT), Knowledge of Features of Learning Mathematics (KFLM), and Knowledge of Mathematics Learning Standards (KMLS).

RESEARCH METHOD

This research adopts a qualitative and narrative approach (Corbin & Strauss, 2008) to characterize the specialized mathematical knowledge mobilized in teaching mathematics to students with visual impairments. Data collection was conducted through video recordings of classroom sessions, given their ability to capture the richness and complexity of classroom interactions for later analysis (Widjaja et al., 2019).

Participants

The sampling strategy was chosen to align with the study's specific context and practical constraints. Three special education teachers working in a school for students with visual impairments (blind or with low vision) were invited to participate. These teachers worked in 1st-, 3rd-, and 6th-grades classrooms, thereby covering a range of

primary education levels. Their professional development course spanned 10 semesters (5 years), with an average of 1.5 courses related to mathematics (Piñeiro & Calle, 2023). Each teacher's classroom comprised 8 students, the maximum allowed in Chilean special schools. All participant teachers had at least 5 years of experience teaching students with visual impairments and were selected because their students did not present additional disabilities associated with their low vision or blindness. Informed consent was obtained from all participants and students' families, ensuring ethical compliance. Data handling adhered to the university's Ethics Committee guidelines, guaranteeing integrity throughout the research process (Ethics Report No. 179/2024).

Data collection and analysis

The data consisted of video recordings of mathematics classes with a focus on natural numbers and arithmetic operations. For this study, a mobile camera with a 360° range was used to report 3 classes, each lasting approximately 80 minutes. The recordings focused on the teacher's interventions with students and the knowledge mobilized by the teacher during these interactions.

Content analysis (Kuckartz, 2019) was applied to the class transcripts, which were segmented into episodes and sub-episodes based on conversational criteria. Special attention was given to teachers' interventions and interactions with students. Units of analysis corresponded to specific recording segments (Krippendorff, 2004) obtained by deductively applying the MTSK analysis categories (domains and subdomains). Thus, a deductive content analysis approach was employed to organize the units of analysis, identifying the categories mobilized by the teachers. Subsequently, an inductive content analysis was conducted to deepen the findings for each MTSK domain. This approach allowed the identification and establishment of specific patterns of knowledge. The analysis centered on the MTSK framework rather than the individual participants.

RESULTS

The results are organized according to the domains and subdomains of the MTSK framework.

Mathematical knowledge

Regarding KoT, in the categories of *definitions*, *properties*, and *principles*, teachers mobilize knowledge of place value, grouping and decomposition, numerical sequence, and certain properties of addition, when asking students questions or explaining some of the steps of the addition algorithm. For example, one teacher asked a student, "Here, right? Where would you place it? In the unit, or in the ten?" thereby mobilizing the role of place value (arranging digits into columns for addition). Teachers also mobilize knowledge of additive calculation *procedures*, such as the standard algorithm and strategies like counting on, counting back, and decomposition. These were evident when teachers prompted students to explain the steps of their calculation strategies. For

instance, one teacher inquired, “Yes, but you counted everything again... you put them together and counted them all again. Did you keep the largest number and add to it?”. On the other hand, regarding systems of *representations*, teachers mobilized knowledge by suggesting concrete representations for students to use or relating different types of representations. Emphasis was placed on concrete representations (cubes, Braille boxes, cubarithm, etc.) and symbolic representations (i.e., Braille code). Pictorial representations adapted for visually impaired students, such as embossed notebooks, were also utilized. For example, a teacher constantly reminded students to write the number sign in Braille before a digit: “Okay, now let’s write the number. Number sign and then the 7. And the number sign?” Finally, on *phenomenology*, the teachers mobilized this knowledge at two levels. First, they use everyday contexts (grocery shopping or use of time) or meaningful contexts (students’ activities) to situate and motivate tasks. For example, one of the teachers provided personalized and contextualized instructions: “Here I have two bags of sugar because I want to make meringue with lots of sugar. I already have eggs at home. But I have two bags of sugar. How much does each bag cost?”. Second, teachers also mobilize knowledge about the types of additive *situations*, particularly interpreting addition as part-whole composition.

Regarding KSM, the teachers mobilized knowledge about *connections* in terms of *simplification* (within the numerical domain) of an *increase in complexity* (regrouping in addition); and *transverse connections* (with other subjects). An example of the latter is presented in the following excerpt, where the teacher links the writing of numbers in the Braille system with the writing of letters in this same system. The digits are represented with the same 4-dots configurations that respectively identify the first 10 letters of the alphabet, with the letter “f” corresponding to the number 6 (distinguished only by the preceding number sign—a special symbol used to indicate that the characters following it should be interpreted as numbers rather than letters).

Student: Now I’m going to do the number sign. Or not yet?

Teacher: Yes, do it.

Student: Number sign. Now I’m going to do the 6. What was the 6 again? One, two, four. One, two, four.

Teacher: Yes, it’s the “f”.

Regarding KPM, teachers also mobilize knowledge of practices in mathematics — related with mathematical language— specifically related to the formal aspects of Braille writing. In this regard, the teachers were emphatic in reminding students the use of number sign. For example, one teacher pointed out to a student: “Yes, there is the number sign, and there is the six, ok? And it has the dots one, two, four”.

Pedagogical content knowledge

Regarding KFLM, teachers exhibited a range of *personal learning theories*. These included beliefs that mathematics is inherently difficult, that speed is an essential

aspect of mathematical learning, and that tasks should be completed step by step. For instance, one teacher addressed a student's error by commenting to the class, "What happened here? Something strikes me about this group. I know that there's a little person here who usually says this is boring, but when it comes to their turn, they're not as quick as I'd expect. It seems like some students don't practice at home at all. Not at all." *Emotional aspects of learning mathematics* were also mobilized, as teachers encouraged students, validated correct responses, and normalized mistakes. For example, when a student missed a mark on the Perkins Braille typewriter, the teacher reassured them, "Good. Go on. Yes! If you missed a... Yes. There! Yes, one mark is missing, but don't worry. Just continue. Keep going."

Regarding KMT, teachers mobilized *personal theoretical knowledge specific to mathematics teaching*, such as the role of the teacher as a guide, the importance of classroom practice, and the existence of a hierarchy in the use of representations. For example, the following excerpt shows how a teacher reminds the student that the activity must be done first with concrete representations and then with pictorial representations: "Now, well, here, quantity first, then move on to the machine. The sequence is: quantity, that is, seeing the set, seeing the number; touching the number, quantity; and then writing it. Got it?". On the mobilized knowledge related to teachers' *strategies*, they diversified their teaching to meet the needs of individual students. This involved adapting numerical domains and using varied concrete materials tailored to each student's characteristics. Diversification was also evident in interaction spaces, as teachers ensured active engagement by asking questions to all students in the classroom. That is, interactions often involved directed questioning, as one teacher asked, "What numerical range are you currently working with?". This differentiated instruction was complemented by a *guided teaching approach*, with teachers mediating step-by-step procedures through explicit verbal instructions or physical guidance, as illustrated by the following excerpt: "Now, look carefully at what you are doing. Use the strategy I taught you. Place one hand on this box and the other on what you're writing. Take a good look. Again. Come on, let's go. Look, look carefully [The teacher holds the student's hands]. That hand has to be in the box on the left. Look carefully at what you're doing. Look, touch. Remember it's one, two, three, four, five, six. Let's go. The same here. One, two, three, four, five, six. Let's go. Try again." Questions were employed to guide step-by-step problem-solving, to shape the above-mentioned strategies or to close lessons with metacognitive prompts, such as, "So, what did we work on today?" Other mobilized strategies were to make explicit what the student should perform, as well as to take the student's hands to guide them in the use of the material that they explored manually. Teachers also redesigned tasks in response to student performance, sometimes repeating tasks to ensure mastery of specific skills. For example, in the following excerpt a teacher asked a student to repeat a well performed exercise: "Ok. Let's write it again. Number 8. Again. Okay, let's go." Then, the knowledge mobilized in relation to *resources and teaching materials* included an awareness of specialized tools for students with visual impairments. These tools

included the Perkins Braille typewriter, embossed notebooks, Braille boxes in different sizes according to the familiarity with the system of each student, nestable cube sets, large-print guides (for students with low vision), and the cubarithm (a tactile matrix for Braille numbers—Manolino et al., in press). For example, when guiding a student, a teacher demonstrates knowledge of using the cubarithm to make additions:

Teacher Look... very good! [She holds the student's hands and guides them on the cubarithm while she speaks]. You placed the 1 here. Then 1 plus 0...

Student ¿1?

Teacher And below is the...

Student ¿9?

Teacher 9... 1 plus 9

Student 10

Teacher Very good. So down here, is it going to be 1 or 0?

Finally, in terms of KMLS, the knowledge mobilized by the teachers was related to the *learning objectives* by explicitly stating what the students needed to achieve during classes. For instance, the technique was through questions as shown in the following excerpt: “What did you work on today in particular? Your classmate already told me she worked on adding rods. What about you?”. Teachers also mobilized knowledge about the *level of procedural development* manifested at the moment of explaining what the students should know. For example, a teacher paraphrases a student's response by pointing out: “So, the goal was to group by tens, right? That's what we've been practicing this week.”

DISCUSSION AND CONCLUSIONS

While prior research has examined the gaps in specialized knowledge among special education teachers, this study highlights the unique knowledge these teachers have developed to meet systemic demands related to mathematics teaching. Thus, this study explores the professional knowledge that three teachers mobilize to meet the demands of the school system, focusing on the complex web of knowledge that can be analyzed through the MTSK framework (Carrillo et al., 2018). Specifically, the study aims to identify the specialized knowledge that special education teachers mobilize when teaching natural numbers and arithmetic operations to students with visual impairments. It is important to clarify that this study does not aim to prescribe what the knowledge of mathematics teachers for students with visual impairments should be. Instead, it seeks to describe the knowledge that has been mobilized in the present context. While this study uses a qualitative approach, the findings build on extensive prior international collaborations with special education teachers, suggesting that the state of affairs described here is not limited to these three teachers but may reflect broader trends in similar educational settings.

A key finding of this study is the alignment between teachers' knowledge and their use of specific materials designed for students with learning disabilities (Gersten et al.,

2009). Teachers used at least three types of representations—concrete, pictorial, and symbolic—and promoted transitions between them (e.g., cubarithm, embossed notebooks, and Braille numbers). Consequently, teachers mobilized knowledge related to representations (KoT) and teaching resources or concrete materials (KMT). These materials and representations were adapted to the characteristics of each student, highlighting the mobilization of knowledge about the characteristics of learning mathematics (KFLM). Another prominent aspect of the mobilized KoT involved procedural knowledge and elements of the decimal number system essential for performing the addition algorithm.

However, these strengths were limited by a focus on answering procedural questions about the steps of the algorithm rather than exploring underlying mathematical ideas. For example, the connections between different representations during these steps or the reasons why the steps work were not addressed. This limitation may stem from the teachers' personal beliefs about teaching (KMT) and learning (KFLM), which could be described as traditional, emphasizing repetition and speed over conceptual understanding.

From our perspective, while these teachers have incorporated different teaching strategies, their application remains situated within the framework of a *traditional* mathematics class. This finding underscores the need for new specialized teacher professional development courses for special education in disciplinary (mathematics) context. This involves viewing mathematics as a teaching and learning space for developing critical and reflective mathematical thinking, rather than merely performing arithmetic procedures.

In conclusion, it is essential to redefine mathematics education for all students, emphasizing not only procedural aspects but also cultural and formative dimensions. This holistic approach would better support the diverse needs of students, including those with learning barriers, and foster a more inclusive and meaningful mathematics education.

Funding and Acknowledgements

Chilean ANID Fondecyt Iniciación 11240835; ANID's FOVI240238; DIUMCE 13-2024 EFA; Italian INdAM (GNSAGA project CUP_E55F22000270001); and the MTSK network, sponsored by the AUIP.

References

- Allsopp, D. H., & Haley, K. C. (2015). A synthesis of research on teacher education, mathematics, and students with learning disabilities. *Learning Disabilities: A Contemporary Journal*, 13(2), 177-206.
- Blömeke, S., Jentsch, A., Ross, N., Kaiser, G., & König, J. (2022). Opening up the black box: Teacher competence, instructional quality, and students' learning progress. *Learning and Instruction*, 79, 101600. <https://doi.org/10.1016/j.learninstruc.2022.101600>

- Brandwand, A., & Johnson, N. (2016). Effective methods for delivering mathematics instruction to students with visual impairments. *Journal of Blindness Innovation and Research*, 6(1). <http://dx.doi.org/10.5241/6-86>
- Carrillo, J., Climent, N., Montes, M., Contreras, L. C., Flores-Medrano, E., Escudero, D., Vasco, D., Rojas, N., Flores, P., Aguilar, Á., Ribeiro, M., & Muñoz-Catalán, M. C. (2018). The mathematics teacher's specialised knowledge (MTSK) model. *Research in Mathematics Education*, 20(3), 236–253. <https://doi.org/10.1080/14794802.2018.1479981>
- Corbin, J., & Strauss, A. (2008). *Basics of qualitative research* (3rd ed.). Sage.
- Firestone, A., Aramburo, C., & Cruz, R. (2021). Special educators' knowledge of high-leverage practices: Construction of a pedagogical content knowledge measure. *Studies in Educational Evaluation*, 70. <https://doi.org/10.1016/j.stueduc.2021.100986>
- Gersten, R., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P., & Flojo, J. (2009). Mathematics instruction for students with learning disabilities: A meta-analysis of instructional components. *Review of Educational Research*, 79(3), 1202–1242. <https://doi.org/10.3102/0034654309334431>
- Klingenberg, O. G., Holkesvik, A. H., & Augestad, L. B. (2019). Research evidence for mathematics education for students with visual impairment: A systematic review. *Cogent Education*, 6(1), 1–19. <https://doi.org/10.1080/2331186X.2019.1626322>
- Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* (2nd ed.). Sage.
- Kuckartz, U. (2019). Qualitative text analysis: A systematic approach. In G. Kaiser & N. Presmeg (Eds.), *Compendium for early career researchers in mathematics education* (pp. 181–198). Springer.
- Lambert, R. (2018). “Indefensible, illogical, and unsupported”; countering deficit mythologies about the potential of students with learning disabilities in mathematics. *Education Sciences*, 8(2), 72. <https://doi.org/10.3390/EDUCSCI8020072>
- Manolino, C., Bernareggi, C., & Armano, T. (in press). Enhancing an early inclusive mathematical learning? Investigating Cubarithm's semiotic potential for children with visual impairments. In *A mathematics education perspective on early mathematics learning* – POEM 2024.
- Piñeiro, J. L., & Calle, J. P. (2023). Mathematical knowledge for teaching in the initial education of special education teachers. *Acta Scientiae*, 25(5), 118-143.
- Tan, P., Padilla, A., Mason, E. N., & Sheldon, J. (2020). *Humanizing disability in mathematics education: Forging New Paths*. NCTM.
- van Garderen, D., Thomas, C. N., Stormont, M., & Lembke, E. S. (2013). An overview of principles for special educators to guide mathematics instruction. *Intervention in School and Clinic*, 48(3), 131–141. <https://doi.org/10.1177/1053451212454006>
- Widjaja, W., Xu, L., & Jobling, W. (2019). Examining primary school teachers' professional noticing through a video-based methodology. In L. Xu, G. Aranda, W. Widjaja, & D. Clarke (Eds.), *Video-based research in education* (pp. 66–82). Routledge.

ANALYSIS OF TEACHERS' EXPERTISE REGARDING COMBINATORICS

Birte Pöhler and Charlott Thomas

University of Potsdam, Germany

Students often have difficulties with combinatorics. In dealing with these teachers play a major role. Therefore, our study aims at providing insights into teachers' expertise regarding combinatorics. For this purpose, eleven teacher interviews were analyzed in terms of content knowledge (CK) and pedagogical content knowledge (PCK) related to the convention of $0!$ and orientations regarding the teaching and learning of combinatorics. The analyses show significant differences in their knowledge about this convention and various teacher orientations, such as perceptions of their own incompetence, students' difficulties, time constraints, and challenges in teaching combinatorics.

THEORETICAL BACKGROUND

As various studies show, students often have difficulties in solving combinatorial tasks (e.g., Lockwood, 2010). Besides the lack of understanding, which is often described as a cause of students' difficulties (e.g., Halani, 2012) content-specific teacher expertise in combinatorics, the focus of this study, could also play a role (Ambrus et al., 2017).

Conceptual framework for content-specific teacher expertise

For conceptualizing teacher expertise for a specific content (like combinatorics) we refer to Prediger's (2019) *Conceptual framework for content-specific expertise*, which is based on the ideas of Bromme's (2014) construct of teacher expertise. It explicitly highlights the content-related aspects of teachers' expertise and combines cognitive and situated perspectives, similar to Bromme's construct. While the latter addresses teachers' classroom practices, which are not the focus here, cognitive perspectives are conceptualized by knowledge (content knowledge (CK), focusing on combinatorics, and pedagogical content knowledge (PCK), focusing on teaching and learning of combinatorics), and orientations (e.g. Kunter et al., 2013). The use of the framework in research aims at investigating teachers' content-specific expertise. This can be useful, for example, to make an informed decision about what teachers need to learn in professional development (PD) courses for the specific topic (here combinatorics) (Prediger, 2019).

The framework of Prediger (2019) consists of *jobs*, representing typical, often complex situational demands of subject-matter teaching relevant to the learning content (e.g. teaching the explanation of $0!$) and *practices*, defined as the recurrent patterns of teachers' utterances and actions for coping with the jobs. Teachers' practices can thereby be characterized by underlying *pedagogical tools* (e.g., tasks or learning

materials), as well as the *categories* and *orientations* that implicitly or explicitly informs their actions.

Since research on teachers' categories and orientations in combinatorics does not yet appear to be very well developed (see the following sections), this study will not examine teachers' practices in combinatorics lessons. Instead, this paper aims to reduce the research gap on combinatorics at teacher level and focus on the analysis of their content-specific categories and orientations.

Possible categories addressed by teachers regarding the concept of $0!$

Categories are conceptual knowledge elements, stem for example, from CK or PCK, that filter and focus teachers' categorical perception and thinking. To create rich, content-related learning opportunities for PD courses, the categories must be specified in content-specific ways (Prediger et al., 2022). This is illustrated below using a specific aspect of combinatorics: the concept of $0!$. The CK includes knowing that $0!$ is defined as 1, as well as the associated explanatory approaches; however, these approaches are also part of PCK, as the teachers need to know how to explain these approaches to students.

While $0! = 1$ is a defined convention, explanations can justify its validity beyond being an abstract rule. Cankoy (2010) identified three approaches for explanations in a study with 58 teachers: *the rule-based*, *the pattern-based*, and *the algebraic-based*. The *rule-based approach*, adopted by about half of the teachers, presents $0! = 1$ as a fixed rule without justification, promoting memorization. The *pattern-based approach* explains $0! = 1$ by extending the factorial rule (e.g., $4! = 24$, $3! = 6$, $2! = 2$, $1! = 1$, so $\frac{1!}{1} = 0!$), supporting structural understanding. This approach was used by a third of experienced teachers but only 3% of inexperienced ones. The *algebraic approach* derives $0! = 1$ from factorial properties (e.g., $\frac{n!}{n} = (n-1)!$, where $\frac{1!}{1} = 0!$), requiring deeper algebraic knowledge. This approach was used by a quarter of experienced teachers and a third of inexperienced teachers. Overall, experienced teachers favored conceptually based approaches (pattern-based and algebraic-based), while inexperienced teachers leaned toward rule-based or algebraic-based explanations (Cankoy, 2010).

Possible teacher orientations on the teaching and learning of combinatorics

Following the conceptual framework for content-specific teacher expertise (Prediger, 2019), orientations refer to content-related and general beliefs that implicitly or explicitly guide the teachers' perceptions and prioritizations of jobs, such as teaching combinatorics in general or a specific concept like $0!$. Possible examples of orientations include beliefs about the content (here combinatorics) or students' learning processes (Schoenfeld, 2010). As mentioned above, little research exists on combinatorics at the teacher level in general or on teacher orientations. One documented orientation in teaching combinatorics (though not explicitly named as

such) is that mathematics teachers often perceive combinatorics as one of the most difficult topics to teach (Annin & Lai, 2010, p.402):

”Mathematics teachers are often asked: What is the most difficult topic for you to teach? Our answer is easy: it's teaching students to count. Although the computations in combinatorics are not necessarily difficult, the concepts can be challenging. Many times, no rigid procedures or formulas can be directly employed to solve the problems, and students simply do not know where and how to begin approaching these problems.”

The impression that combinatorics is challenging for many students - as expressed at the end of the quote - is shared by Hungarian teachers, who were asked about their attitudes, the usefulness, and difficulties on combinatorics (Ambrus et al., 2017). Additionally, most teachers expressed a positive attitude toward combinatorics and found it useful, even beyond mathematics.

RESEARCH QUESTION AND METHODOLOGICAL APPROACH

This study aims at analyzing teachers' expertise in combinatorics, particularly regarding addressed categories and expressed orientations. Therefore, this paper focuses on the following two research questions:

RQ1: What content knowledge and pedagogical content knowledge do teachers show in interviews regarding the explanation of $0! = 1$?

RQ2: What orientations in general and towards the teaching and learning of combinatorics do teachers express in interviews?

To answer the research questions, the data was used from semi-structured interviews lasting approximately 30 minutes. For RQ1 (addressed categories), teachers were asked to what extent they had introduced the convention $0! = 1$ and how they would respond if a student inquired why this equality holds. For RQ 2 (addressed teachers' orientations), relevant questions addressed their attitudes toward combinatorics, the relevance of teaching the topic, and typical challenges by teaching and learning it.

A total of eleven teachers participated in the interview. All teachers work at secondary schools in Germany and were teaching students (aged between 17 and 18) in their final year before finishing school. According to the teachers, combinatorics had been taught in the last six months. This means that the combinatorial operations (variation, permutation and combination each with and without repetition) were taught in four teaching blocks, as outlined in the German curriculum.

All interviews were fully transcribed and analyzed using a qualitative content analysis (Mayring, 2015). To address RQ1 and analyze the categories within teachers' content-specific expertise, several categories were used to describe the explanation of $0! = 1$: The categories C1 to C3 were adapted from Cankoy (2010) and extended by category C4 to include a verbal combinatorial approach. Additionally, category C5 was added because, unlike Cankoy (2010), this study did not specify the result $0!$ (Table 1). The content knowledge (CK) includes here the knowledge of the definition $0! = 1$, while pedagogical content knowledge (PCK) encompasses the explanatory approaches, since

the teachers were asked explicitly how they would explain this definition to their students and not how they would explain it to themselves.

Category	Explanation of the approach
C1: Algebraic approach	$0! = 1$ is derived from the algebraic characteristics of the factorial.
C2: Pattern approach	It is explained that factorials follow a pattern.
C3: It is a rule approach	It is described that $0! = 1$ is a fixed rule.
C4: Verbal combinatorial approach	It is argued with words, there is only one way to arrange nothing.
C5: Wrong or no complete explanation	A wrong/no result is stated or only the beginnings of a justification are given.

Table 1: Deductive (C1-C3) and inductive (C4-C5) categories for explaining $0! = 1$

To address RQ2, passages were extracted from the interview transcripts to reconstruct the teachers' orientations towards combinatorics. Categories were derived inductively from the extracted passages. Thereby three main categories emerged, enabling the reconstruction of orientations regarding teachers' own perception of combinatorics, the teaching of it and the relevance of combinatorics and their perception of learning it by the students (see similar in Schoenfeld, 2010). Figure 2 illustrates the specific categories assigned to the above-mentioned main categories based on the data analysis. The inductively developed category system facilitates a systematic analysis of teachers' orientations toward the teaching and learning of combinatorics. For instance, it identifies which orientations are more prevalent or considered ambivalent.

Category	Description
Own teacher's perception of combinatorics (OP)	
AOP-F	Affective perspective: Highlights the fun and positive attitudes towards combinatorics.
AOP-F	Affective perspective: Includes negative attitudes towards combinatorics such as lack of interest or fun
COP-C	Cognitive perspective: Reflects teachers' own challenges regarding combinatorics & solving combinatorial tasks
Own teacher's perception of teaching combinatorics (OP-T)	
ACOP-T	Teaching perspective: Teaching combinatorics is perceived as particularly difficult
COP-T	Curricular Perspective: Combinatorics is perceived as having limited relevance for final examinations (like abitur)
TOP-T	Time perspective: Limited time planned or available for addressing the topic in class.
Own teacher's perception of the relevance of combinatorics (ROP-S)	
ROP-S	Student-centered perspective: Teachers see an importance of combinatorics for students' future learning
ROP-S	Student-centered perspective: Teachers see no relevance of combinatorics for student's further learning
Teacher's perception of learning combinatorics by the students (SP)	
ASP-F	Affective perspective: Teachers emphasize that stochastic topics are fun for students
CSP-C	Cognitive perspective: Teachers report students' competencies regarding combinatorics
CSP-C	Cognitive perspective: Teachers report that students often have difficulties by solving combinatorial tasks
CSP-A	Cognitive perspective: Teachers see an ambivalence between students' expertise in combinatorics & other mathematical topics

Fig. 1: Inductive categories for teachers' orientations regarding combinatorics

An independent person coded a total of 100% for RQ1 and 25% for RQ2 of the data set. The intercoder reliability rate κ was calculated using MAXQDA 2024 and was

$\kappa = .80$ for the category system of RQ1 and $\kappa = .72$ for the category system of RQ2. Thus, it can be classified as almost perfect for RQ1 and substantial for RQ2 (Landis & Koch, 1977). Consensus was subsequently reached for the divergent cases.

RESULTS

To address the question of what content-specific expertise teachers demonstrate regarding combinatorics exemplarily, we will first describe their reconstructed content (CK) and pedagogical content knowledge (PCK) of $0!$. Then, we will present their orientations toward combinatorics. Regarding RQ 1, which analyzed the *content knowledge* of $0!$, seven out of eleven teachers knew that $0!$ is defined as 1 (C1-C4). The other four stated, for example, “I wouldn't even have known that $0!$ is 1. That's what my table in the formulary is for. I can't say anything about it, because you would say 0. Factorial is zero times nothing.” In addition to stating that the value of $0!$ is zero, some teacher described it as “undefined” (C5). These responses revealed gaps in CK. Regarding RQ 1, for the *pedagogical content knowledge* of $0!$, two teachers used the algebraic approach (C1), arguing that the definition is logical to maintain the connection between permutation and arrangement without repetition. One teacher explained that when n is equal to k , the permutation formula $n!$ can be derived from the arrangement formula $\frac{n!}{(n-k)!}$, making it necessary to define $0!$ as 1. This teacher also mentioned the neutral element of multiplication as an alternative explanation but preferred the algebraic approach for its clarity for students. This illustrates the connection between the CK (“ $0!$ is 1 because of the neutral element of multiplication”) and the PCK. The teacher would possibly address the CK in the advanced course, but would rather favor the algebraic approach, as this is closer to the students. One teacher argued with words, explaining that there is only one way not to order anything (C4). Four teachers stated that $0! = 1$ by definition (C3). They offer no further justification even when we asked how they would address students' questions about it. For example, one said, “I just tell them $0!$ is defined as 1.” Notably, no teacher used the pattern-based approach, which justifies $0! = 1$ by following the factorial sequence.

To address RQ2, which analyzed teachers' *orientations* toward combinatorics, the eleven interviews were examined using the inductively developed category system. Figure 2 provides an overview of the results. It should be noted that the categories ROP-S (student centered perspective: no relevance of combinatorics for further learning; T1, turn 12), ASP-F (affective perspective: combinatorics provide fun for students; T12, turn 86) and CSP-C (cognitive perspective: experiencing competence of students regarding combinatorics; T14, turn 22) only appear once and are therefore not listed in figure 2. The interviews reveal ambivalent orientations of the eleven teachers regarding their affective attitude towards combinatorics (categories AOP-F and AOP-E). Whereas six teachers expressed a rather positive attitude (like T11, turn 8: “Yes, combinatorics is actually a really nice topic.”), one teacher saw positive and negative aspects (T10, turn 6: “When it comes to combinatorics in teaching, I am ambivalent.”)

and three teachers said that they did not enjoy it (like T1, turn 24: “If you know you didn't like it at school and at university, you have to overcome that to do it again.”).

Teacher (T)	Own teacher's perception of combinatorics			Own teacher's perception of teaching combinatorics			Own teacher's perception of the relevance	Own teacher's perception of learning combinatorics by students	
	AOP-F	AOP- F	COP- G	ACOP-T	COP-T	TOP-T	ROP-S	CSP- G	CSP-A
1		#24	#24, 30	#12, 14, 24	#14			#16, 24, 26	
2			#20, 82, 84	#2, 6	#10, 24	#24, 86			
3	#4,6		#2, 4, 8, 10, 69			#54		#10, 12, 14	#12
4	#6				#8	#78	#8	#16, 18, 20, 22	
5	#8					#14, 95		#26	
6		#4, 10	#6, 16, 22, 69	#4, 6	#8, 74	#10, 22		#6	#6
7		#4	#6, 9, 10	#4	#12, 14	#43		#4, 8	
8	#8		#90				#10	#26, 30	
9	#4				#6	#6, 12	#6	#4, 10	
10	#8	#6	#4	#84		#4, 86			
11	#8		#18, 52	#8	#8	#8, 22		#8	#18

Fig. 2 Teachers' orientations reconstructed from the interviews

Except for three teachers, all teachers expressed their own experience of incompetence regarding combinatorics (COP-~~G~~). Especially, the following aspects were perceived as difficult: reading / understanding the tasks (T6), dealing with the technical terms (like permutation or variation; T2), formulating own tasks (T2, T6), identifying / differentiating the combinatorial figures (T2, T3, T7) and justifying them (T10). Just under half of the teachers describe combinatorics as difficult to teach (ACOP-T). T1 explained it as follows: “You know you didn't like it. [...] That was the biggest challenge to sell it to the students anyway that it's fun and you can do a lot with it.” Most teachers also expressed that combinatorics considered to be of low relevance from a curricular perspective (COP-T), as it is rarely integrated in final examinations. This was noted for example by T9 (turn 6): “It is often not to be expected that many questions on combinatorics are integrated in the abitur and that is my focus to prepare the students well for it”. Furthermore, there is little time for the teaching and learning of combinatorics and especially for practising it by the students (TOP-T), as eight teachers remark. For example, T6 (turn 10) expressed: “I'll be honest with you, the treatment of combinatorics is not really long”. Nevertheless, some teachers describe the topic as being relevant to students (ROP-S), for example as a tool for further learning in probability (T4, T9), but also because it contributes to foster logical competences (T8) and has an everyday relevance (T10). Almost all teachers express students' challenges (CSP-~~G~~) with learning combinatorics. Especially, identifying and differentiating the combinatorial figures (T1, T3, T4, T5, T7, T9, T10), reading and understanding the tasks (T1, T3, T4, T5, T6, T7), dealing with technical terms (T7), making references to probability theory (T3), dealing with more complex tasks (T4, T8, T11) and formulae with variables (T4) were described as difficult. In addition,

three teachers describe a similar observation. Teacher 6 (turn 6) formulates the observed phenomenon as follows: “Combinatorics is often particularly challenging for students who excel in topics such as function theory or vector algebra. In the past, I've noticed again and again that those who have difficulties there are suddenly good at probability, combinatorics and statistics. That is a bit strange.”

DISCUSSION AND CONCLUSION

Both the analysis of CK and PCK of teachers regarding $0!$ revealed substantial gaps in the knowledge of four out of eleven teachers. This suggests that some teachers rely heavily on external aids, such as tables or the textbook, rather than their own conceptual knowledge. This is particularly problematic, as it prevents students from developing a conceptual understanding due to the teacher's lack of knowledge. Some of the teachers may not (or cannot) convey that this convention has meaningful justifications. In summary, these results underscore the need for targeted professional development to address gaps in both CK and PCK. Strengthening teachers' understanding of $0!$ and providing them with various explanatory strategies can enable them to foster a deeper student understanding of combinatorial concepts. This is especially important, as a solid foundation in combinatorics supports broader mathematical reasoning and problem-solving skills (Lockwood et al., 2020). The need for teachers' professional development in combinatorics can also be derived from the analysis of teachers' orientations. Most teachers expressed content-specific incompetencies. In addition, some teachers perceived teaching combinatorics as difficult (also documented by Ambrus et al., 2017 for Hungarian teachers). This could influence their teaching as the topic being addressed rather briefly and in little depth in class, which in turn can cause students' difficulties. The fact that some teachers show rather negative emotions regarding combinatorics (in contrast to Ambrus et al., 2017 results) underlines that in a PD course on combinatorics not only CK and PCK should be taught, but also the orientations should be addressed and further developed in the longer term. Regarding the generalizability of the study results, it should be noted that the sample size is limited as data are only available from eleven interviews. An increase in the sample size could therefore be beneficial in the future. Furthermore, the analysis of the categories addressed by the teachers only covers one central concept of combinatorics - $0!$. In addition, it must be considered that we used self-reported data, which may show distortions (e.g. due to socially desirable answers). These data can only be used to reconstruct selected components of teacher expertise - in this case, addressed categories and expressed orientations. To be able to draw a more comprehensive picture of teachers' expertise, it would be useful to observe teachers' lessons on combinatorics and derive typical teacher practices from these. In addition to the already mentioned and necessary implications for practice, namely that high-quality content-specific PD courses for teachers need to be developed and researched, the study also reveals research desiderata. For example, other aspects of combinatorics (like combinatorial operations) and their teaching or connections between teacher expertise and students' competences could be investigated.

References

- Annin, S. & Lai, K. (2010). Common Errors in Counting Problems. *Mathematics Teacher*, 103(6), 402–409.
- Ambrus, G., Rácz, N. K., & Vancsó, Ö. (2017). How Hungarian teachers think about combinatorics and its teaching. *Scientific Issues of Jan Długosz University in Częstochowa. Mathematics*, 22, 11–22.
- Bromme, R. (2014). Der Lehrer als Experte: *Zur Psychologie des professionellen Wissens (Vol. 7)*. Waxmann Verlag.
- Cankoy, O. (2010). Mathematics Teachers' Topic-Specific Pedagogical Content Knowledge in the Context of Teaching a^0 , $0!$ and $a \div 0$. *Educational Sciences: Theory and Practice*, 10(2), 749–769.
- Halani, A. (2012b). Students' ways of thinking about enumerative combinatorics solution sets: The odometer category. *Proceedings of the 15th Annual Conference on Research in Undergraduate Mathematics Education*, 1, 231–245.
- Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S., & Neubrand, M. (2013). *Cognitive activation in the mathematics classroom and professional competence of teachers. Results from the COACTIV Project*. Springer.
- Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1), 159.
- Lockwood, E. (2010). An investigation of Post-Secondary Students' Understanding of Two Fundamental Counting Principles. *Proceedings of the 13th Annual Conference on Research in Undergraduate Mathematics Education*.
- Lockwood, E., Wasserman, N. H., & Tillema, E. S. (2020). A case for combinatorics: A research commentary. *The Journal of Mathematical Behavior*, 59, 100783.
- Mayring, P. (2015). Qualitative Content Analysis: Theoretical Background and Procedures. In A. Bikner-Ahsbahr, C. Knipping, & N. Presmeg (Eds.), *Approaches to Qualitative Research in Mathematics Education: Examples of Methodology and Methods* (pp. 365–380). Springer.
- Prediger, S. (2019). Investigating and promoting teachers' expertise for language-responsive mathematics teaching. *Mathematics Education Research Journal*, 31, 367–392.
- Prediger, S., Roesken-Winter, B., Stahnke, R. & Pöhler, B. (2022). Conceptualizing content-related PD facilitator expertise. *Journal of Mathematics Teacher Education*, 25(4), 403–428.
- Schoenfeld, A. H. (2010). *How we think: A theory of goal-oriented decision making and its educational applications*. Routledge.

MAPPING ROUTINES FOR ANALYZING TUTOR-STUDENTS INTERACTIONS

Cristina Poli, Giulia Lisarelli, Anna Baccaglini-Frank

University of Pisa

This study investigates interactions between a tutor and a pair of students during mathematics activities involving digital artifacts designed for low-achieving high school students. Through two cases, we explore the potential of a mapping representation with respect to gaining insight into the dynamics of the discursive routines in tutor-students interactions.

INTRODUCTION

This study is part of a broader research project, DynaMat, aimed at designing and implementing sequences of activities with digital artifacts. The activities consist of a digital environment and an assignment related to it. The project studies mathematics learning of low achieving students, and how to overcome their difficulties. Indeed, the literature suggests that digital artifacts, that offer interactive and dynamic interactions, can serve as valuable tools for enabling mathematics learning (e.g., Sinclair & Yurita, 2008; Antonini et al., 2020). However, the role of the teacher (or in our case, tutor) is fundamental, in interacting both with the students and with the digital artifact (e.g., Mariotti, 2009). Moreover, even when tutors have access to previously designed and experimented materials, they still have to make various moves on-the-fly during the implementation. They also may reiterate patterns of interaction with the students that they are not warned about by the accompanying guidelines.

In light of these considerations, we are interested in analyzing tutor-students discursive interactions in activities with digital artifacts, looking for interesting (with respect to students' learning) phenomena and patterns. Through the analysis of two cases, using a mapping representation, we will pursue such an intent, using a discursive approach. Specifically, we explore the potential of our mapping representation tool.

THEORETICAL FRAMEWORK

In this study, we adopt a participationist approach, that emphasizes joint participation in shared cultural activities, focuses on social, predominantly linguistic interaction, and studies learning, as it occurs through interaction (Sfard & Cobb, 2022). Within this perspective, the theory of commognition conceptualizes learning mathematics as becoming a participant in mathematical discourses, and the routinization of learners' actions is considered the process through which it may occur. *Routines* are a central construct: “patterned ways in which mathematical tasks are being performed” (Sfard, 2012, p.2). Lavie et al. (2019) define a *task situation* (TS) as any situation in which an individual feels the need to act. This includes, for example, situations set by requests

of a tutor to students. When faced with a given TS, different students may respond in different ways that, also, do not necessarily match the tutor's expectations. This is because the student interprets the TS based on what is available in their *precedent search space*, the set of past situations sufficiently similar to the current one and therefore recalled. The *task* is what the student interprets as needing to be done in the specific situation. The *procedure* is how the student believes they should act. The pattern of actions resulting from the combination of the interpreted task and the adopted procedure to solve it is called a *routine*. Each routine is therefore highly personal.

To identify interesting phenomena and patterns in the tutor-students interactions, one interesting aspect to capture is, for a given TS, mismatches between the routine expected by the tutor and the routine enacted by the student. We will call the tutor's task the *intended task*, since in our project tutors were aware of the activities' design principles. Such a mismatch may also occur on sub-routines (Sfard, 2023) that a tutor might try and hint at to support the students. For the tutor these sub-routines would be *bonded* (each routine feeds into another, that is, the outcome of the first routine serves as an input to the second), but not necessarily for the students. The digital artifact will also have a role in our scenario, that we will try to capture in relation to the expected and enacted routines (Baccaglini-Frank et al., in press).

These are the sort of phenomena and patterns that we are interested in trying to map in the two episodes in focus.

Research objective

Our research objective is to study the tutor-students interactions in the context of activities with digital artifacts. Specifically, we seek to identify mismatches in each TS between the intended routines and sub-routines and those implemented by the students, and other interactional patterns perpetuated by different tutor-students triads when working on a same TS.

METHOD

Cora, Letizia, Bene, and Fra volunteered as participants in the study, recognizing themselves (this was confirmed by their teachers and an interview) as low-achieving (in mathematics) students in 10th grade (15 years old). They participated, in pairs, to an intervention comprising a sequence of activities on functions carried out over 5 2-hour meetings; each pair worked with a different tutor (both researchers on the DynaMat project). During the activity sessions, the students had one tablet showing the digital artifact, on which the screen and audio recording were activated. Each session was also recorded by a fixed camera.

Description of the activity in the digital environment

The Moving Arrows Environment (MAE) presents a red arrow, labelled a , on a horizontally oriented number line; the endpoint of the arrow is marked with a red tick mark and is directly draggable bound to the oriented number line. Parallel to the red

arrow, there are four black arrows labelled a , $-a$, 1 , and -1 . The endpoints of the black arrows a and $-a$ move indirectly, one opposite to the other, as the red tick mark is dragged, maintaining the same length of the red arrow, while 1 and -1 are fixed arrows. The “Add arrow” button allows the construction of new arrows by creating different consecutive copies of the black arrows on a horizontal line. The endpoint of the last copied arrow is marked with a black tick mark that moves in correspondence with the dragging of the red tick mark. Figure 1 shows the result of the construction of a copy $-a$ followed by two consecutive copies of -1 .

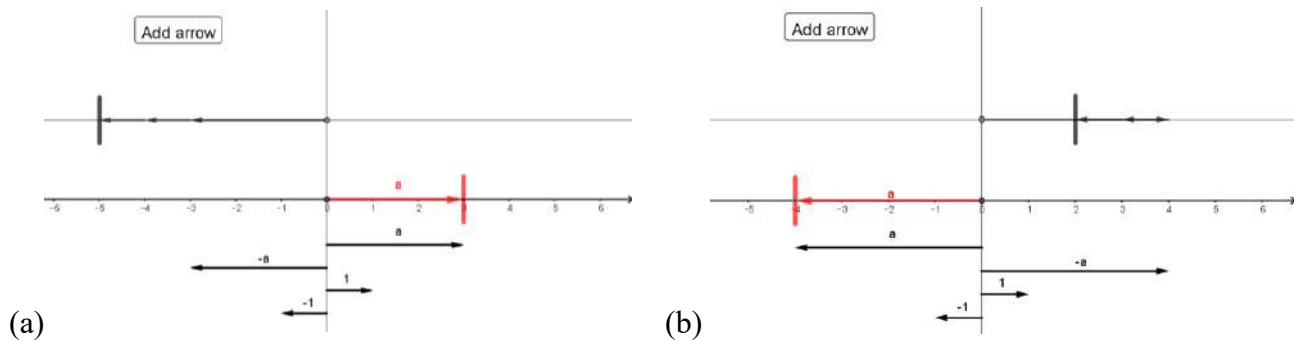


Figure 1: Two screenshots of the MAE in which a copy of $-a$ and two consecutive copies of -1 have been created with the ‘Add arrow’ button, and $a=3$ (a); $a=-4$ (b).

According to the guidelines that were given to the tutors in the DynaMat project, the initial TS they provided to students was: use the “Add arrow” button to make a construction corresponding to “the opposite of a , minus 2” (which can be obtained, for instance, as in Fig. 1). Here the tutors did not have a rigid protocol to follow; the next TS in the sequence asked to predict (i.e., to answer without dragging the red arrow) on which value the black tick mark will go if the red tick mark is dragged onto some given values. After every prediction, the students would be asked to drag and check their answer.

Mapping the routines

The analytic tool we devised to map the mathematical routine is an adaptation to our needs of what other commognitive researchers have called Discourse Mapping Diagram (DMD) (e.g., Heyd-Metzuyanim et al., accepted; Weingarden et al., 2019). The DMD allows to graphically represent the routine enacted by the student, as well as the sub-routines hierarchically nested in it. It is structured so that the main TS lies at its root. The routines are represented by blue-gray boxes in which the blue part contains a task, and the gray part contains a narrative or a procedure. The task interpreted by the students is reconstructed based on the enacted procedure or on the students’ narratives. The number denoting each box-pair marks the hierarchical enumeration of the sub-routine.

In this study, we adapted this tool to include the tutor’s interactions in the mapping. More specifically, we added red-gray boxes in which the red part contains the TS as given by the tutor and the gray part contains the intended procedure (Fig. 2). These

boxes are aligned in the DMD so that the task interpreted by the students is under the TS assigned by the tutor. Moreover, the boxes are connected to each other by: a horizontal continuous arrow if there is a mathematical bond between routines; a horizontal dotted arrow if there is a bond but one of the routines is outside the context of the TS (that in our case means out of the MAE); a vertical continuous arrow if there is a sub-routine (and it could be also oriented upwards, for example, when the students' procedure in the sub-routine feeds the TS given by the tutor).

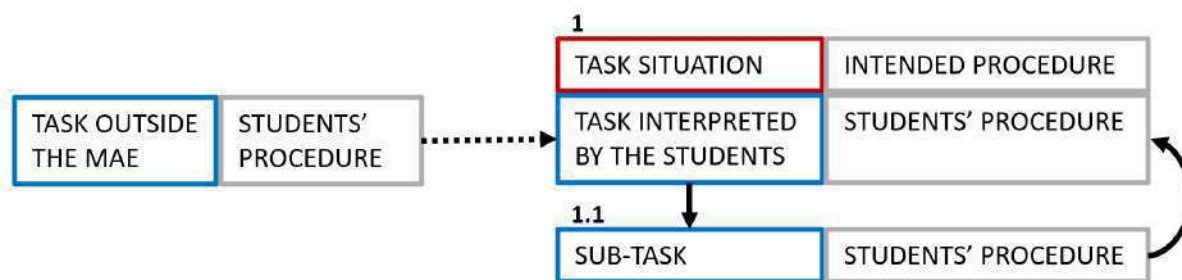


Figure 2: Example of tutor-students Discourse Mapping Diagram

DATA ANALYSIS

Figure 3 and Figure 4 depict the DMDs for the cases of Bene and Fra and of Cora and Letizia as they work with the respective tutors on the task situation described above.

The case of Bene and Fra

In the first part of the episode the tutor intervenes only to provide the TS (1) and then lets the students work independently, in line with the activity guidelines. We identified 3 sub-tasks, as Bene and Fra enact the following procedures for constructing “the opposite of a , minus 2”: they drag the red tick mark onto 2 (1.1), they construct “the opposite of a ” by adding a copy of $-a$ (1.2), they construct “minus 2” by adding two times $+1$ (1.3). The tutor asks them how they did the construction (2). We note a mismatch between the students’ procedure and the intended one. In front of this, the tutor launches a new TS, which she does not make explicit initially (empty box 3); instead, she breaks it down into explicit sub-tasks: the first (3.1) and the second (3.2) refer to the current position of the red tick mark, i.e. the value 2, and the tutor asks what the result of $-2-2$ should be. Bene and Fra answer 4, instead of -4 , as the tutor expected, and they ask for clarification on the given task.

In response, the tutor takes up the initial TS (3.3) and asks them to write in symbols the expression “opposite of a minus two” (3.3.1). The students then write $-a-2$, as expected by the tutor. This also leads the students to recognizing a mistake. The first mismatch (1.3) now seems to have been overcome. The tutor (4) recalls the first construction made by the two students in the MAE (corresponding to $-a+2$) asking “What did you do by adding two arrows “1”?”. The students’ narrative differs from the intended one: Bene and Fra say that the arrow represented $-a+a$, since in the specific case $a=2$, and so it was the “same thing [as $-a+2$]”. To face this mismatch, the tutor

formulates two sub-tasks (4.1 and 4.2) that focus on the symbolic writing and not on the construction within the MAE: she asks the students to write in an algebraic form the previous construction (4.1). Different scriptures emerge: $-a+a$, $-a+1+1$, $-a+2$ and $-2+1+1$. Not expecting all of these narratives, the tutor asks for the differences between the scriptures (4.2). A discursive pattern emerges: the tutor articulates the given request in sub-tasks (4.2.1; 4.2.2; 4.2.3), that focus on the validity of the scriptures “if a was 3”, to which the students respond authoring intended narratives.

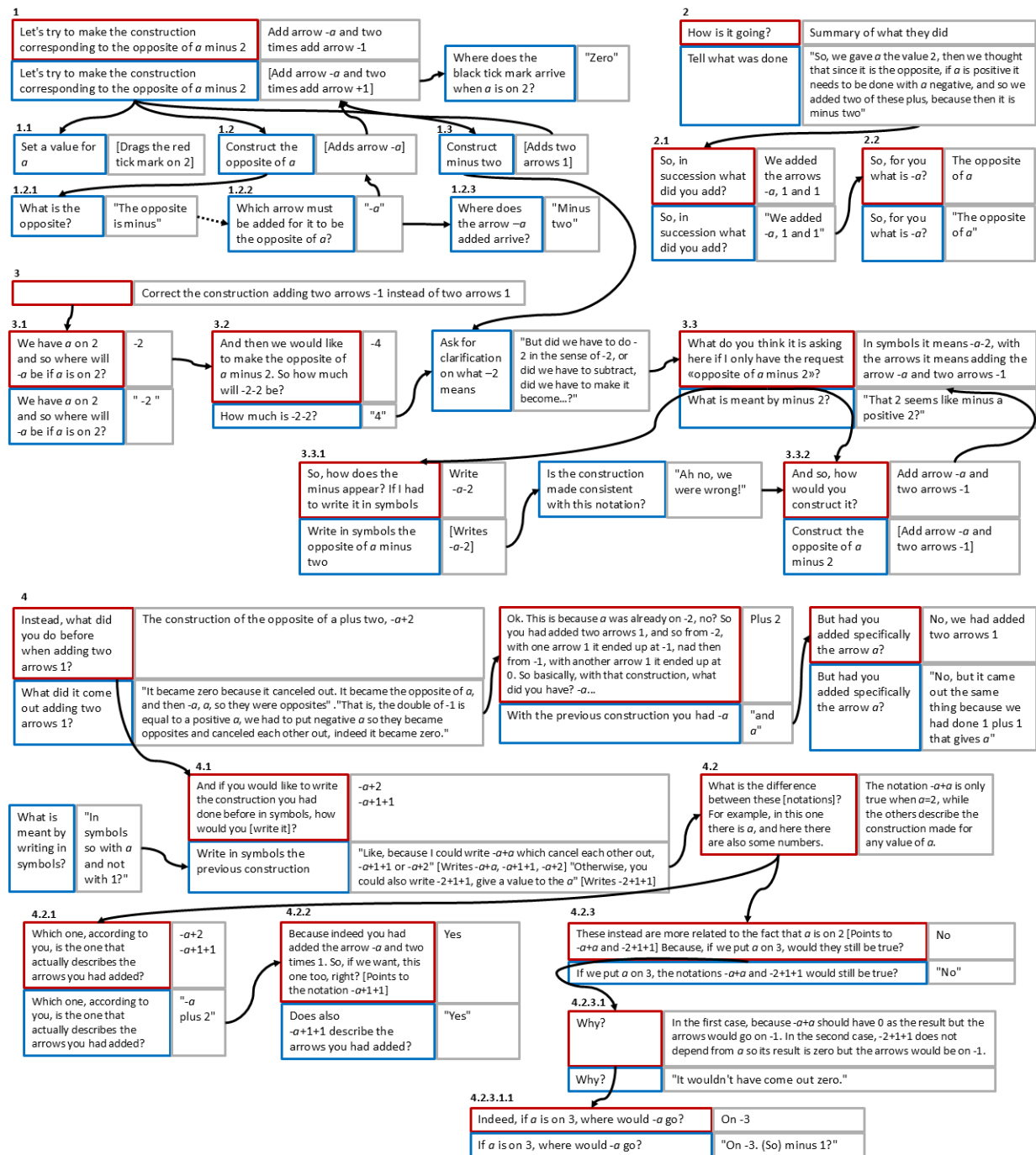


Figure 3: Bene and Fra DMD

The case of Cora and Letizia

In the first part of the episode the tutor (a different one) intervenes only to describe the TS (1) and then lets the students work independently. Cora and Letizia perform four sub-routines without interacting with the tutor: in 1.1, they write the expression $b = -a - 2$; in 1.2, they set a value to drag the red tick mark; in 1.3, they construct the opposite of a by adding the arrow $-a$; and in 1.4, they construct -2 by adding two arrows -1 . Therefore, the students' procedure is the same as the intended one. Then the tutor introduces a second TS (2) by asking whether the construction aligns with what they have written and why.

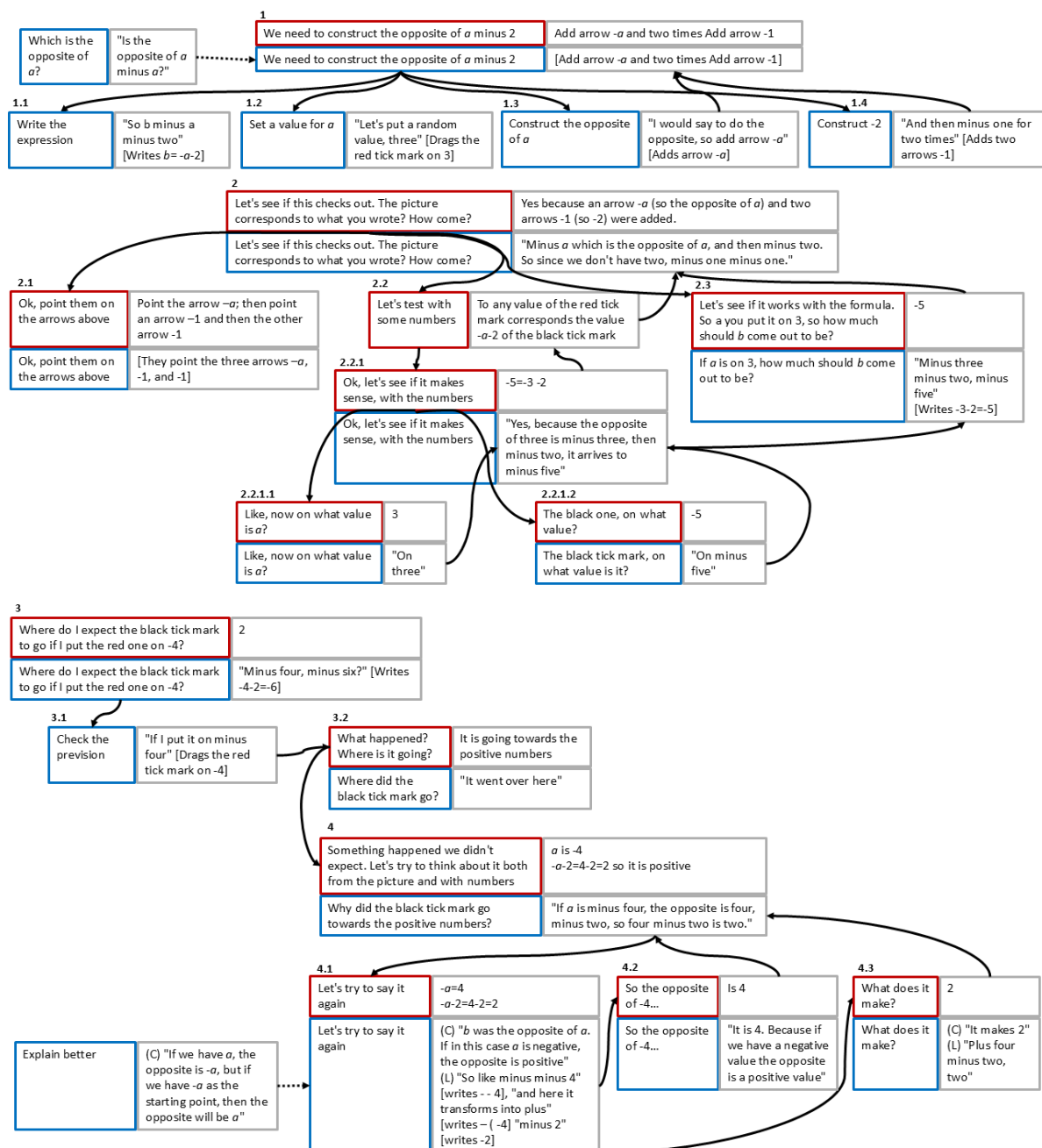


Figure 4: Cora and Letizia DMD

Then, he articulates it into sub-tasks, by asking the students to indicate the arrows corresponding to the written expressions (2.1) and to check using numerical values (2.2), such as the number 3, where the red tick mark is positioned at the time. The students' routines also in this case are bonded and lead them to answer correctly.

The next TS (3) in the designed sequence asks Cora and Letizia to make a prediction about which value the black tick mark will go onto, if the red tick mark is dragged on -4. Here there is a mismatch between the intended procedure and the students' one, because the students write -6 as their prediction. Cora and Letizia then drag a to -4 and receive feedback from the MAE which does not match with their prediction (3.2). This opens the final TS posed by the tutor (4), who asks to explain why it happened, encouraging them to "think about it both from the picture and on the numbers". The task interpreted by the students seems to focus on numbers and on the symbolic writing that Cora exploits recalling the endorsed narrative "if a is negative, the opposite is positive" and Letizia modifies by adding a parenthesis to isolate -4 (for further details on this episode see also Baccaglini-Frank et al, in press).

DISCUSSION AND CONCLUSION

The analyses of the two episodes shed light onto patterns in the interaction between each tutor and their students, showing how these can vary significantly, even for a same TS and when the tutors are researchers on the same project! In both cases, we identified instances of mismatch between the intended routine and the one implemented by the students. In response to these mismatches, the two tutors react differently: in the case of Bene and Fra, the tutor does not explicitly ask for the construction to be corrected, but instead she breaks the original intended task into sub-tasks. Instead, in the case of Cora and Letizia, the tutor focuses on the feedback provided by the MAE, and then he asks the students to explain what happened, allowing them to interact with each other to reach an agreement.

The analysis of the mapped routines also reveals that the digital artifact played a different role in the two case studies. In the first case, there is little interaction with the MAE. Indeed, the discourse is only about the result of the construction in the MAE. In the second case, we identify a different phenomenon: the discourse relies on the interaction (and the feedback received) with the MAE. One hypothesis to be further explored is that such a phenomenon is related to the tutors' different familiarity with the activities and the MAE.

In conclusion, DMDs appear to be a valuable analytical tool not only for mapping students' routines (Heyd-Metzuyanim et al., accepted) but also, with our adaptation that overlaps the tutor's routines, for gaining insight in fine-grained detail into the tutor-students interactions and into the role of the tutor. Finally, although the context of this study is very specific (a remedial intervention with low-achieving students) and we cannot assume generalizability of our adaptation of the DMD tool, we believe that it yields the potential to be insightful also for analyzing teacher-students discursive

interactions in the regular mathematics classroom. Through further research we hope to investigate further in this direction.

Acknowledgment

This study is part of the funded project DynaMat (PRIN 2020BKWEXR); experimental work conducted at CARME (www.carme.center), UNISER Pistoia Srl, Italy. We also acknowledge the MIUR Excellence Department Project awarded to the Department of Mathematics, University of Pisa, CUP I57G22000700001.

References

- Antonini, S., Baccaglini-Frank, A., & Lisarelli, G. (2020). From Experiences in a Dynamic Environment to Written Narratives on Functions. *Digital Experiences in Mathematics Education*, 6, 1–29. <https://doi.org/10.1007/s40751-019-00054-3>
- Baccaglini-Frank, A., Bonadiman, C., Lisarelli, G., & Nannini, B. (in press). Cora and Letizia's attempt to construct a discourse about the expression "-a". *Proceedings of the 15th International Congress on Mathematical Education*.
- Heyd-Metzuyanim, E., Cohen, B., Elbaum-Cohen, A., & Tabach, M. (accepted). The inextricable nature of regulative and mathematical routines in the process of de-ritualization. *Educational Studies in Mathematics*.
- Lavie, I., Steiner, A., & Sfard, A. (2019). Routines we live by: From ritual to exploration. *Educational Studies in Mathematics*, 101(2), 153–176. <https://doi.org/10.1007/s10649-018-9817-4>
- Mariotti, M.A. (2009). Artifacts and signs after a Vygotskian perspective: the role of the teacher. *ZDM Mathematics Education*, 41, 427–440. <https://doi.org/10.1007/s11858-009-0199-z>
- Sfard, A. (2012). Introduction: Developing mathematical discourse—Some insights from communicational research. *International Journal of Educational Research*, 51–52(3), 1–9. <https://doi.org/10.1016/j.ijer.2011.12.013>
- Sfard, A. (2023). The devil's finest trick: Routines that make teachers matter against their better judgement. *Journal of Curriculum Studies*, 55(1), 21–36. <https://doi.org/10.1080/00220272.2023.2172359>
- Sfard, A., & Cobb, P. (2022). Research in Mathematics Education: What Can It Teach Us about Human Learning? In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 467–485). Cambridge University Press. <https://doi.org/10.1017/9781108888295.029>
- Sinclair, N., & Yurita, V. (2008). To be or to become: How dynamic geometry changes discourse. *Research in Mathematics Education*, 10(2), 135–150. <https://doi.org/10.1080/14794800802233670>
- Weingarden, M., Heyd-Metzuyanim, E., & Nachlieli, T. (2019). The realization tree assessment tool – examining explorative participation in mathematics lessons. *Journal of Mathematical Behavior*, 56. <https://doi.org/10.1016/j.jmathb.2019.100717>

ANALYZING INQUIRY-BASED LEARNING WITH PROBLEM POSING IN MATHEMATICS TEACHER EDUCATION: INSIGHTS FROM THE EXPANSIVE LEARNING CYCLE

Gresa Pozhegu-Ermeni¹, Svein Arne Sikko¹ and Benedikte Grimeland¹

Norwegian University of Science and Technology (NTNU), Norway

We employ qualitative methods to report on an empirical study where we investigated the implementation of inquiry-based learning (IBL) with problem posing (PP) through pre-service teachers' (PSTs') experience in two communities: on-campus and in school-practice. Our theoretical framework is based on Engeström's activity theory, and the data were analyzed using content analysis. Guided by expansive learning cycle, we examined the processes of questioning, reflection, and adaptation of IBL with PP in different communities. The study contributes valuable insights into how to overcome the challenges of integrating IBL with PP, leading to more effective and sustainable inquiry-based teaching practices in mathematics education. Future research should refine IBL with PP teaching practice and explore their long-term impact.

INTRODUCTION AND THEORETICAL BACKGROUND

In this research report we explored how problem posing (PP) activities can enhance and develop inquiry-based learning (IBL) approaches and its implementation. This was achieved through pre-service teachers (PSTs) as they learned and developed an IBL lesson plan with PP in two communities: on-campus and in school-practice. IBL is a student-centered approach of teaching and learning that engages students in exploring problems, developing solutions and explaining ideas (Engeln et al., 2013), and has the potential to promote confidence, persistence, independence, and creative thinking in mathematics education (Laursen et al., 2016). Despite IBL's widely recognized benefits, it's often hindered by tensions between teachers and students, arising from its weak implementation that disrupts the classroom environment (Leat et al., 2014). Implementing IBL in teaching practices faces challenges such as motivating students, designing confidence-building investigations, ensuring background knowledge, providing sufficient time, and meeting learning context constraints (Edelson et al., 1999, Potari et al., 2024). Overcoming these challenges can offer students valuable opportunities to better understand mathematics and make the teaching easier for teachers, but there are still few comprehensive studies of IBL implementation and impact on mathematics education (Dorier & Maass, 2020). Thus, the gap between the positive aspects of IBL and difficulties in implementing it in the classroom need to be bridged. A method to support IBL implementation and enhance inquiry can be the use of PP in the IBL lesson design. PP refers to an activity of creating new problems to investigate a particular situation as well as reformulating known mathematics problems (Lavy & Shriki, 2007). Research (e.g. Zhang & Cai 2021) has also pointed out that there is a lack of studies that analyze the content and features of teaching through PP,

and there is an interest in integrating PP into classrooms. We see it as important to study how PP can be included as a part of IBL to facilitate PSTs' learning and teaching, and this can help to inform research to support teachers of mathematics in learning how to integrate IBL in their teaching.

In this paper we analyzed the implementation process of IBL with PP, starting from the knowledge that PSTs acquired at the campus community to the experiences that PSTs gained in school-practice. To do this, we drew on activity theory as our theoretical framework, as it emphasizes that learning occurs through social and cultural interactions with a focus on understanding how people interact with tools, rules and communities (Engeström, 1987). Engeström (1987) created a model to understand human activities, which includes six nodes: subjects, tools, goals, community, rules, and division of labor and this model is often shown as a triangle. This study adopted the third generation of CHAT by comprising two interacting activity systems (see Figure 1) to investigate the interaction between the activity systems of the PSTs in two communities: on-campus and in school-practice.

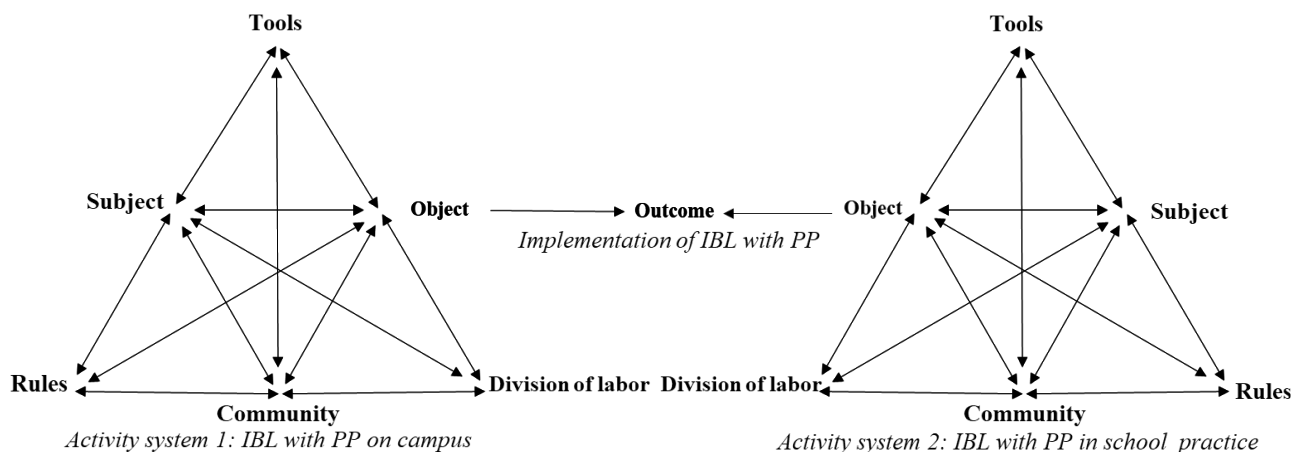


Figure 1. The activity system adapted from Engeström, (1987)

An important principle of CHAT are contradictions, which serve as a central source of change and progress (Engeström, 1987). In an activity system, contradictions can build up over time and lead to changes and these contradictions can occur within a single activity system or between different activity systems (Engeström, 2010). By consciously bringing contradictions into the activity system, one or more nodes can be changed, eventually transforming the whole system as a result. Foot (2014) explains that within an activity system, contradictions are not seen as failures or deficits, but rather as starting points for further development. There are four types of contradictions: 1) primary contradictions that occur within a single node of an activity system when subjects encounter conflicting values or demands at that node; 2) secondary contradictions that occur between different nodes of the activity system when a new element is introduced and must be integrated into the existing activity; 3) tertiary contradictions that occur between nodes when participants need to adopt more advanced methods to achieve the object of the activity; and 4) quaternary

contradictions that occur between neighbouring activity systems when changes in one activity system create conflicts with another. Understanding contradictions within the system is crucial for data analysis, as contradictions can drive expansive learning by prompting new objects and motives when individuals address conflicts (Engeström & Sannino, 2010). Expansive learning refers to a process where the participants go beyond the initial task to create new knowledge and practices and involves a sequence of learning actions aimed at resolving such contradictions. Observing and navigating the contradictions that drive learning and shape outcomes within an activity system can help us understand and foster expansive learning transformations. The logic of the expansive cycle is that it begins when an existing activity system is questioned and ends when a new pattern of activity has been consolidated (Larsen et al., 2022). In this way, the expansive learning cycle consists of seven basic actions (see Figure 2): 1) participants start by questioning the current practice, recognizing that it isn't working well due to internal contradictions or gaps; 2) they then analyze the situation in detail, using historical and empirical analysis to understand it's root causes; 3) participants propose and develop a new model to resolve these contradictions, envisioning a more advanced approach; 4) they examine and elaborate on the new model, experimenting with it to see how it works in practice and if it addresses the identified problems; 5) if the new model shows promise, they implement it, adapting and fine-tuning as needed; 6) participants reflect on and evaluate the effectiveness of the process, discussing and critiquing the changes to gain a deeper understanding and identify any new contradictions that arise; 7) finally, they consolidate the new practice into a stable form, though new contradictions may emerge, starting the learning cycle again.

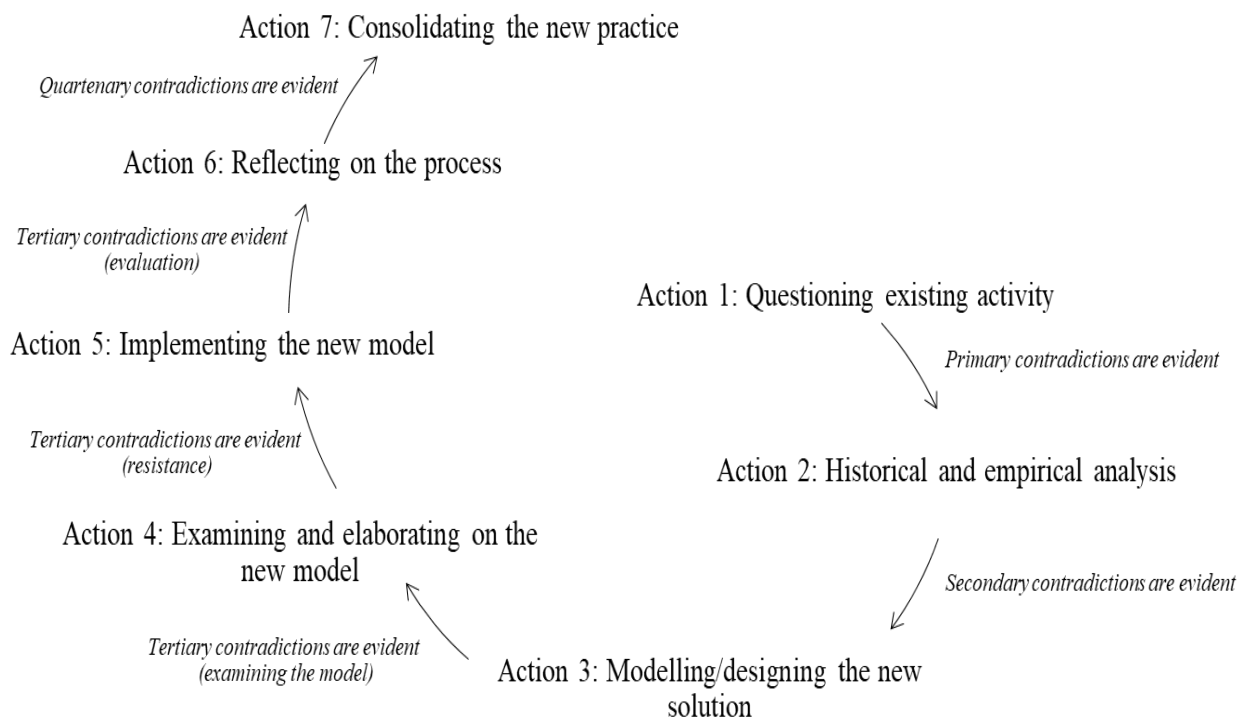


Figure 2. Expansive learning cycle: sequence of actions and contradictions

CHAT emphasizes that learning is not just about acquiring existing knowledge but about creating new knowledge and transforming practices to adapt to changing environments (Engeström, 2016). Based on this, we examined the interaction between the two activity systems which share a common outcome: implementation of IBL with PP through PSTs' experience. Accordingly, our research question is: How do contradictions encountered during the cycle of expansive learning drive inquiry-based learning with problem posing?

METHODOLOGY

This study employed an empirical qualitative research methodology to investigate the implementation of IBL with PP during PSTs' campus and school-practice, examined through the lens of activity theory. Activity theory provided a framework to understand contradictions within and the dynamics between the two activity systems, guiding the formulation of the research question and the analysis of the implementation process. Using the expansive learning cycle, we identified and modelled actions, clarifying the process of forming and resolving internal contradictions within these activity systems. We first identified the actions and the subjects participating in them, then matched the contradictions to the expansive learning cycle model. In our analysis, we did not cover all possible contradictions within or between the nodes of a single activity system or across two activity systems. Instead, we concentrated on four specific types of contradictions, identifying nodes involved in the learning process. It is crucial to keep in mind that each contradiction occurred as part of the activity system as a whole entity.

DATA COLLECTION AND METHODS

The participants of this study were three teacher educators who were also the researchers and authors of this study, and 80 Norwegian PSTs for grades 5–10, all in their second semester of the first year. The first author was both a researcher and a teacher of the PSTs. The PSTs were informed about the study and gave their written consent to participate in the research. To gain a comprehensive understanding of the PSTs' expansive learning cycle with IBL and PP, we employed diverse data collection methods. Data were collected through a combination of qualitative methods, including thirty-six lesson plans, sixteen reflective diaries, ten interviews, and seventeen practice reports, and were analyzed using content analysis. The PSTs engaged with IBL and PP through lectures and coursework on campus. On campus, we gathered their initial lesson plans and reflective diaries. The PSTs participated in school-practice and implemented an IBL lesson plan including PP, aligned with their practice school's agenda. As part of their coursework, PST were asked to write a practice report, which were included as data. During the school-practice, the PSTs were responsible for planning, structuring, and leading group discussions. After the school-practice, we conducted semi-structured in-person interviews, each lasting 30–60 minutes. The interviews focused on the PSTs' challenges and attitudes toward implementation of IBL with PP in both campus and school communities. English was not the participants' first language, but as the interviewer did not master Norwegian to a sufficiently fluent

level at the time, the interviews were conducted in English, recorded, and transcribed. The multi-faceted approach to data collection ensured a rich dataset, enabling a thorough exploration of the PSTs' experiences with IBL and PP. Participation was voluntary and could be withdrawn at any time without providing a reason, and participants were ensured confidentiality and anonymity.

RESULTS: ANALYSIS OF ACTIONS AND CONTRADICTIONS IN ACTIVITY SYSTEMS

We outlined each action along with the contradictions that occurred through the cycle of expansive learning. Such an approach provided a comprehensive view of the challenges and experiences of implementing IBL with PP on campus and during school-practice. Actions 1 to 4 took place on campus. In the first action, researchers questioned the current practice of teaching math using the IBL approach. After reviewing the literature on IBL, researchers decided to introduce PP to enhance its effectiveness. Researchers aimed to identify how PSTs could use PP as a tool to enhance inquiry in an IBL lesson plan. This led to a primary contradiction in the object node, highlighting the need for new teaching practices within IBL. Consequently, this involved including new tools (such as PP) to enhance IBL teaching practices in classrooms. In the second action, PSTs created lesson plans incorporating IBL. Researchers reviewed these plans to determine whether the PSTs were able to incorporate IBL-approaches. It was found that the implementation was less apparent than anticipated, with several plans lacking key elements of the inquiry-based approach. This indicated that the PSTs faced challenges in integrating the inquiry process into their lesson plans. In the third action, researchers developed an advanced approach by introducing PP for PSTs to include in their lesson plans, with the aim to see if inquiry would enhance with the inclusion of PP. PSTs subsequently designed and developed a new lesson plan incorporating PP with IBL. The analysis of the revised lesson plans showed that PP activities offered prompts that guided and supported PSTs through the IBL process. This led to a secondary contradiction: the new tools, like using PP to enhance inquiry, contradicted with the PSTs' original lesson plans, which lacked IBL and were teacher-centered. In the fourth action, researchers examined and developed a practice-task for PSTs, asking them to implement IBL in their school-practice teaching, including PP. The fifth action took place in school where PSTs wrote their lesson plan and implemented it with their students. After school-practice the PSTs wrote a report detailing the implementation process of IBL with PP in school. In the sixth action, researchers analyzed the practice reports, including the final lesson plans, and interviewed PSTs about the challenges of implementing IBL with PP. PSTs reflected on the entire process, assessing the challenges and outcomes. Researchers aimed to address these challenges and identify which PP tools supported the learning objectives in an IBL approach. The reflection process, provided researchers with a deeper understanding of the difficulties in implementing IBL with PP, leading to a tertiary contradiction: using IBL with PP as an advanced teaching method conflicted with PSTs' actions and interactions in school-practice. The contradiction made it

difficult for PSTs to implement IBL with PP, along with other challenges like managing a whole school-class on their own:

PST: Well, yeah, I think it was quite nice, but I did one of the lessons alone and then we were few adults, so it was a bit harder to actually see what they were doing and guide everyone, but it was easy when we were like six

After implementation and reflection, researchers better understood the challenges and aimed to consolidate the new teaching practice of IBL with PP into an accepted model within the activity system. Quaternary contradictions between actions six and seven, involved reflecting on and evaluating the implementation's successes and failures, as well as the broader learning journey. As per the seventh action, this model based on Engeström's expansive learning cycle, sparked new contradictions during its implementation between campus and school communities, suggesting the need to potentially restart the cycle to ensure full acceptance.

DISCUSSION AND CONCLUDING REMARKS

Through our analysis, we clarified how the implementation of IBL with PP was driven by contradictions encountered during the expansive learning cycle between different activity systems. The primary contradiction occurred on campus activity system and highlighted the need for new teaching practices within IBL in mathematics education, but did not respond directly to a need felt by the PSTs neither on campus nor in school-practice. This fact may be meaningful to mathematics teacher educators, but it does not necessarily affect teachers, who are ultimately responsible for implementing classroom changes. This aligns with the findings of Larsen et al. (2022), who noted that although researchers and teachers occupy different roles in research development, it is unclear how collaboration between them can be achieved. Both studies emphasize the challenge to bridge this gap, so future studies could focus on collaborative inquiry-research that involve both researchers and teachers from the outset. This approach could foster a more integrated development process, ensuring that both perspectives are considered and utilized effectively. Laursen et al. (2016) suggest that IBL is an approach well suited to mathematics departments seeking to strengthen their pre-service teacher preparation offerings in ways consistent with research-based recommendations. Therefore, it is crucial to provide ongoing professional development opportunities through pre- and in- service teachers to enhance their understanding of implementing IBL. The secondary contradiction highlighted the need for straightforward and accessible teaching methods to implement IBL effectively. By making changes to PSTs' teaching practice, researchers contributed to an expanding learning circle showing how teaching should be conducted to enhance inquiry, integrating PP in IBL. A tertiary contradiction was revealed by signs of resistance from PSTs during school-practice. In CHAT, a tertiary contradiction arises when a new, more advanced form of activity is introduced, conflicting with the existing, dominant form (Engeström, 2016). In the context of using IBL with PP, this advanced teaching method represented the new form of activity. Our study found that the teaching practice

of IBL with PP contradicted the PSTs' established routines and expectations in school-practice. This resistance was manifested as difficulty in implementing IBL with PP, as PSTs struggled to adapt to the new demands and responsibilities in school-practice. This is consistent with Potari et al. (2024) findings, which indicated that both teachers and students faced difficulties of engaging inquiry due to differing expectations and modes of working in school and other courses, leading to the creation of tensions. Both studies highlight the challenges and tensions that arise when introducing new teaching practices that differ from established routines and expectations. To address the tertiary contradiction, it is essential to integrate IBL with PP more effectively in mathematics education, ensuring that PSTs acquire adequate knowledge in campus courses to avoid tensions during school-practice. This study suggests several practical steps. First, professional development programs should use the study's findings to improve teaching methods. Additionally, creating collaborative communities where PSTs, teachers, and researchers can share their experiences with IBL and PP will foster a supportive environment. Introducing IBL with PP gradually will help PSTs adapt and integrate it smoothly into their teaching practice. Finally, encouraging PSTs to regularly reflect on their teaching, identify areas for improvement, and develop better IBL strategies will lead to continuous professional growth. Furthermore, the quaternary contradictions between the sixth and seventh actions indicated that while initial steps may have been successful, new challenges arose when the model was applied in different contexts, such as between campus and school communities. These contradictions are a natural part of the expansive learning process (Engeström, 2016), as they reveal areas where the current teaching practices and new approaches clash, necessitating further refinement and adaptation. Referring to the research question, the contradictions encountered during the cycle of expansive learning highlighted opportunities for consolidating teaching practices into future frameworks for implementing IBL with PP. We recommended to encourage continuous reflection and evaluation, foster collaboration between campus and school communities, and be prepared to restart the learning cycle as much as is needed. By embracing these recommendations, the study contributes valuable insights into navigating the challenges of implementing new teaching models, leading to more effective and sustainable IBL teaching practices in mathematics education.

References

- Dorier, J.-L., & Maass, K. (2020). Inquiry-based mathematics education. In B. Sriraman, M. Graven, Y. Shimizu, E. Ruahama, M. Artigue, & E. Jablonka (Eds.). *Encyclopedia of mathematics education* (pp. 384–388). Springer Cham. https://doi.org/10.1007/978-3-030-15789-0_176
- Edelson, D. C., Gordin, D. N., & Pea, R. D. (1999). Addressing the challenges of inquiry-based learning through technology and curriculum design. *The Journal of the Learning Sciences*, 8(3-4), 391–450. <https://doi.org/10.1080/10508406.1999.9672075>

- Engeln, K., Euler, M., & Maass, K. (2013). Inquiry-based learning in mathematics and science: A comparative baseline study of teachers' beliefs and practices across 12 European countries. *ZDM Mathematics Education*, 45(6), 823–836. <https://doi.org/10.1007/s11858-013-0507-5>
- Engeström, Y. (1987). *Learning by expanding: An activity theoretical approach to developmental research*. Cambridge University Press.
- Engeström, Y. (2016). *Studies in expansive learning: Learning what is not yet there*. Cambridge University Press.
- Engeström, Y., & Sannino, A. (2010). Studies of expansive learning: Foundations, findings and future challenges. *Educational Research Review*, 1–24. <https://doi.org/10.1016/j.edurev.2009.12.002>
- Foot, K. A. (2014). Cultural-historical activity theory: Exploring a theory to inform practice and research. *Journal of Human Behavior in the Social Environment*, 24(3), 329–347. <https://doi.org/10.1080/10911359.2013.831011>
- Larsen, D. M., Østergaard, C. H., & Rasmussen, K. (2022). Contradicting activity systems—Learning from large scale interventions that fail to change mathematics teaching practice as intended. *Journal of Mathematics Teacher Education*, 27, 5–28. <https://doi.org/10.1007/s10857-022-09548-1>
- Laursen, S. L., Hassi, M.-L., & Hough, S. (2016). Implementation and outcomes of inquiry-based learning in mathematics content courses for pre-service teachers. *International Journal of Mathematical Education in Science and Technology*, 42(2), 256–275. <https://doi.org/10.1080/0020739X.2015.1068390>
- Lavy, I., & Shriki, A. (2007). Problem posing as a means for developing mathematical knowledge of prospective teachers. *Proceedings of the 31st international conference on the Psychology of Mathematics Education (PME)*, 3, pp. 129–136. Seoul, South Korea.
- Leat, D., Thomas, U., & Reid, A. (2014). Reframing relationships between teachers, students and curriculum – The phenomenon of ‘hybridisation’ in IBL. In P. Blessinger, & J. M. Carfora (Eds.), *Inquiry-based learning for faculty and institutional development: A conceptual and practical resource for educators* (Vol. 1, pp. 101–119). Emerald Group Publishing.
- Potari, D., Metaxas, N., Jaworski, B., & Zachariades, T. (2024). Inquiry mathematics teaching in a university bridging course: challenges for students and teachers. *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education (PME)*, 4, pp. 17–24. Auckland, New Zealand: PME.
- Zhang, H., & Cai, J. (2021). Teaching mathematics through problem posing: Insights from an analysis of teaching cases. *ZDM Mathematics Education*, 53, 961–973. <https://doi.org/10.1007/s11858-021-01260-3>

EYE MOVEMENTS AND SPONTANEOUS FOCUSING ON NUMEROSITY AMONG FIVE-YEAR-OLD CHILDREN

Satu Rajalakso¹, Ilona Södervik¹, Minna Hannula-Sormunen², and

Markku S. Hannula¹

¹University of Helsinki, ²University of Turku

Spontaneous focusing on numerosity (SFON) refers to extent to which children spontaneously recognize and use numerosity in action. SFON is known to be a significant predictor of mathematical skills development. This study is the first one to implement eye-tracking to study SFON. The aim of the present research is to investigate the relations between young children's SFON and their eye movement behaviour during memory card game. The results did not reveal significant differences between how many times participant had fixated on objects they did or did not enumerate. However, the extensive variation in the number of fixations on the enumerated objects suggests distinct enumeration strategies between subitizing- and counting-based enumeration.

INTRODUCTION

Early numeracy skills indicate later mathematical performance. Numeracy at kindergarten predicts arithmetical skills both at the first years of primary school (Aunio & Niemivirta, 2010) but also years later (Hannula-Sormunen et al., 2015; Jordan et al., 2009). One significant factor in development of early numeracy is the extent to which a child spontaneously attends to mathematical aspects during everyday activities (Batchelor et al., 2015; Hannula & Lehtinen, 2005; Torbeyns et al., 2017). Spontaneous focusing on numerosity (SFON) refers to the child's tendency to spontaneously focus on numerosity and use numerical knowledge in action (Hannula & Lehtinen, 2005).

The current research investigates the relations between young children's SFON and eye movement during memory card game. We explore the variation in eye movement behaviour when the child does or does not spontaneously recognize and produce numerosity. The present study is presumably the first one to exploit eye-tracking methods in studying young children's spontaneous mathematical behaviour and the first one to implement this methodology in SFON research.

Our research questions are: 1. How does SFON relate to eye movements during memory card game? and 2. How did eye movement behaviour differ when the object was enumerated, only named, or not named at all?

THEORETICAL FRAMEWORK

Although all children have innate mathematical abilities, studies have shown that there is an extensive variation on how frequently children use these skills. Subitizing-based

enumeration enables exact and immediate numerosity recognition (Trick & Pylyshyn, 1994). Already infants have been found to recognize small quantities (1-3), long before learning verbal counting skills. Counting occurs when quantities exceed the limits of subitizing (Trick & Pylyshyn, 1994). However, to utilise enumeration, enumeration processes must be intentionally activated, and attention focused on the aspect of numerosity (Hannula & Lehtinen, 2005). Spontaneous focusing on numerosity (SFON) refers to the extent to which children spontaneously recognize numerosity in their everyday environment and use them in action (Hannula & Lehtinen, 2005). Previous research has distinguished SFON as a separate mental process from enumeration, though being crucial to exploit enumeration skills (Hannula & Lehtinen, 2005; Nanu et al., 2018).

Children with high SFON tendency frequently recognize numbers of objects and events in everyday surroundings and learn to use that knowledge in action. Thus, spontaneous, i.e., self-initiated recognition of exact numbers of items provides advantageous practice in quantifying skills (Hannula & Lehtinen, 2005; Torbeyns 2017). However, SFON is not stable but develops thorough childhood in a bidirectional relationship with other early numeracy skills and emerges contextually (Elliott et al., 2022; Hannula & Lehtinen, 2005). SFON is proved to be strongly connected to both concurrent and later mathematical skills (e.g., Batchelor et al., 2015; Gloor et al., 2021; Hannula et al., 2010; Torbeyns 2017).

Eye movement behaviour consist of extremely rapid eye movements, saccades, and short times between the saccades, called fixations, when the eye remains stable (Rayner, 1998). Due to saccadic suppression, all new visual information is acquired during fixations (Matin, 1974). Multiple studies over the years have validated the role of eye-tracking methodology when studying cognitive processes and visual attention (Just & Carpenter, 1980). Thus, eye-tracking has been used quite extensively in mathematics education. Important results have been obtained concerning enumeration and pattern recognition (e.g. Baumanns et al., 2024; Schindler et al., 2020; Sprenger & Benz, 2020). Those explorative studies have identified distinct strategies participants use to perceive structures, recognize patterns, and enumerate sets with various quantities and arrangements. However, all previous eye-tracking research of mathematical cognitions is based on tasks that explicitly guide to approach it mathematically. Thus, the aim of the present research is to expand the scope of the research field by investigating children's spontaneous activities and to add our limited knowledge about implementing eye-tracking procedures with young children.

RESEARCH DESIGN AND METHODS

The data collection for an individual child included a single 5–10-minute memory card game further described below. During the session, participants eye movements were tracked for short periods at a time. The test settings were also video, and audio recorded to ensure capturing all verbal and non-verbal utterances. The testing took place in a separate room in participants' own kindergartens on a one-to-one basis with the

researcher (first author). The procedure was designed as short-term, fun, and motivating game to minimize the child's mental and physical burden. Researcher had visited the kindergartens for one day in advance to get to know the children and familiarise with the available space. Hence, children felt more comfortable conducting the task alone with the researcher and interacted openly during the game. During the testing, the researcher and the participant sat next to each other in front of the eye-tracking apparatus. A video camera was placed facing the participant.

The memory card game was an adaptation of the task first presented by Nanu et al. (2020). A detailed description of the game can be found at <https://osf.io/tk8wb>. The present game was adapted as an investigation game where the participants were asked to play as detectives and conduct investigations for their client (the researcher). During the game, the participant was presented cards displayed on the eye-tracking monitor for 5 seconds each during which the apparatus tracked participant's eye movements. The instruction for the participant was to scrutinize the picture for the 5-second period it was displayed on the monitor and immediately after report back to the researcher what they had seen. Participant was asked to provide as detailed description as they could remember so that the "client" can find the matching card from her set of four cards. When the child ended the description, they were asked: "Do you remember anything else?". The trial ended with the researcher playfully finding the matching card regardless of how profound the description was. In the beginning of each trial the researcher quickly flashed her cards for the participant to underline how similar the pictures were. The game consisted of four trials. Each trial contained one participant's card and four researcher's cards, from which one was the same as the participant's and the other three highly similar. The cards were same as in the original task containing sets of 1-19 familiar objects for children which were 1. Lego bricks, 2. art supplies, 3. food items, and 4. train toys. The number of similar items in a single card were small to facilitate subitizing based enumeration. However, the cards included unlimited possibilities to produce any quantities.

We used a remote Tobii Pro Spectrum eye-tracking system, which consists of a 23.8" screen with resolution of 1920x1080 pixels and screen-mounted eye-tracker. The tracker captures data at a sampling rate of 600 Hz with precision of 0.08° and accuracy of 0.17°. Stimulus was presented with Tobii Pro Lab -software. The distance between the tracker and participant was approximately 60cm. No chin rest was used to keep the research situation as natural as possible. The eye-tracker was calibrated using 5-point calibration procedure before the game. Before each picture, a fixation-cross was displayed in the middle of the screen for one second.

Participants

The final data includes 22 native Finnish speaking children (12 female, 10 male) aged 5-6 years (M: 5 years 8 months, SD: 4 months) with normal or corrected-to-normal vision. Seven participants were excluded due to their first language being other than Finnish, deficient language skills, inaccurate calibration of the gaze-tracker, or other

complications during the game. Participants were recruited from two kindergartens in Southern Finland located in average SES neighbourhoods. Written informed consent was obtained from all parents, and children were asked for verbal assent. No indications of mathematical nature of the study were provided to the guardians or the participants to ensure the measurement of spontaneous actions.

Analysis

All participants' responses were transcribed from the video and audio recordings. A SFON score was determined for all participants based on the memory card game. The scoring followed the guidelines in Nanu et. al. (2020). The participant was determined to express SFON if their description included any expressions of exact enumeration, or the coder observed any indications of counting, regardless of their accuracy. For every card, participant was scored 1 if their descriptions included any mentions of numbers of items, or indications of enumeration, and 0 if they did not. Maximum of the task was 4.

The overall quality of the eye-tracking data was good with sufficient accuracy and relatively small amount of data-loss. One trial from three participants contained excessive number of lost fixations and was therefore excluded. The data loss resulted mainly due to unfavourable head postures, eye-openness, or participant turning gaze direction outside of the screen. The final data consisted of 85 trials which all contained at least four seconds of fixation data. The measures used from eye-tracking data were average fixation count and average fixation duration.

The verbatim transcriptions of all recorded descriptions were also further analysed. For each participant and each card, we identified the objects participant had mentioned and whether there were any signs of enumeration in their description. Then, every object in each card was categorised in one of three groups, based on whether the participant had 1.) enumerated ("two yellow blocks"), 2.) only named ("a jar with pencils") or 3.) not named (but fixated on) the object. Then, each object was defined as a separate area of interest (AOI) using Tobii ProLab software to determine how many times participant had fixated on it. An average fixation count was then determined for every participant on enumerated, named, and not named objects.

Skilling-Mack test was used to explore possible variation between three groups. The test was selected because it tolerates data with arbitrary missing observations (Chatfield & Mander, 2009). Wilcoxon Signed Rank tests were conducted as post-tests to further investigate pairwise connections. Non-parametric tests were required since no variable distribution or sampling distribution were normal.

RESULTS

First, we explored whether SFON was related to children's fixation durations and fixation counts. The second focus was in whether children's eye movements were different for those objects that they enumerated in comparison to others.

How does SFON relate to eye movements during memory card game? (RQ1)

To gain an overview on the phenomenon, we compared the participants' SFON to their overall average fixation durations and average fixation counts. We examined the correlations between the participants' SFON scores ($M: 1.73$, $SD: 1.12$, $Mdn: 2.00$) and their average fixation duration and average fixation count first for each card using Spearman's rank correlation coefficient test. Then we calculated average fixation metrics over all four trials and conducted the same correlation comparisons. These tests did not produce any statistically significant results suggesting that SFON does not correlate with average fixation duration or average fixation count.

How did eye movement behaviour differ when the object was enumerated, only named, or not named at all? (RQ2)

For the second research question, a more specific analysis was conducted by comparing the differences between three formed groups: 1.) enumerated, 2.) only named, and 3.) not named with average fixation counts. The analyses were first conducted to each card separately and then to all four cards combined. The descriptive statistics are presented in Table 1. It is relevant to note here, that the standard deviation of fixation counts was notably larger for most enumerated objects than for other objects.

Card	Enumerated				Named				Not Named			
	N	M	SD	Mdn	N	M	SD	Mdn	N	M	SD	Mdn
Lego	10	3.71	2.90	2.25	20	3.31	1.84	3.17	21	1.92	1.17	1.50
Art supply	12	1.97	0.60	2.00	18	1.91	0.61	2.00	21	2.09	0.71	2.00
Food	2	5.75	2.47	5.75	20	3.32	1.26	3.45	17	2.63	0.93	2.50
Train	10	3.26	2.09	2.92	19	2.42	1.54	2.00	19	2.40	0.90	2.25
All	17	3.10	1.69	2.38	22	2.72	0.66	2.70	21	2.22	0.66	2.00

Table 1: Number of participants, means, standard deviations, and medians of average fixation counts by groups in each card and within all cards

Skillings-Mack test demonstrated significant variance between three groups in Lego card ($SM=8.60$, $p=.01$) and in average of all four cards ($SM=6.05$, $p=.05$). No significant results were obtained in art supply card ($SM=0.58$, $p=.75$) or train card ($SM=1.24$, $p=.54$). The test was not possible to conduct with food card since there were only two observations in group 1 including enumerated objects.

Wilcoxon signed-rank test did not reveal any significant difference between enumerated and only named groups in Lego card ($z=-.07$, $r=.01$), art supply card ($z=-.91$, $r=.17$), or train card ($z=-.06$, $r=.01$). The results were similar when testing the averages of all four cards ($z=-.54$, $r=.09$). The test results followed a similar notion also between enumerated and not named groups. The variation remained modest in all cards, Lego ($z=-1.40$, $r=.25$), art supply ($z=-.47$, $r=.08$), and train ($z=-1.13$, $r=.21$). The results indicated a slightly more notable difference between these groups when compared with the average of all cards yet remaining only almost statistically significant ($z=-1.87$, $p=.06$, $r=.30$). Lastly, comparisons between named and not named groups were conducted. These tests reached statistical significance in Lego card ($z=-2.31$, $p=.02$, $r=.36$) and in average of all cards ($z=-2.52$, $p=.012$, $r=.38$). No significant variation was obtained in art supply ($z=-1.04$, $r=.17$), food ($z=-1.04$, $r=.17$), or train cards ($z=-.34$, $r=.05$).

Average fixation counts of all cards in groups enumerated, named, and not named are presented by means and standard deviations in Figure 1.

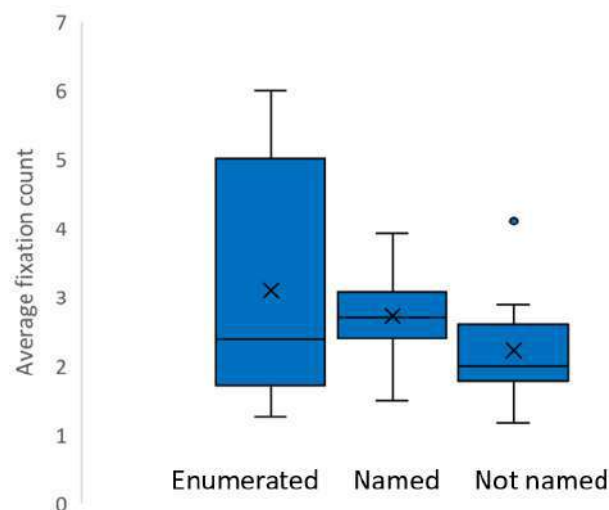


Figure 1: Mean values and standard deviations of fixation count of all cards in groups enumerated, named, and not named. The cross marks the mean value of the group.

DISCUSSION

The present study was the first one to examine the relations between children's spontaneous focusing on numerosity and eye movements, providing valuable new knowledge for both research fields. Very few previous studies have tested children under school age with eye-tracking methods. The current study offers an example for successfully applying eye-tracking with young children.

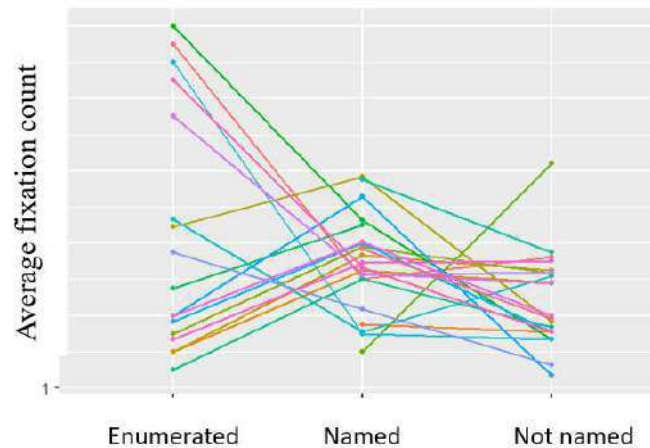


Figure 2: Each participant's average fixation counts of all cards in groups enumerated, named, and not named

The results show that the overall fixation count was not significantly different between objects that were enumerated and those that were named but not enumerated. However, the variation of fixation count was significantly larger for enumerated objects (see Figure 2). Some participants had more fixations, and some fewer fixations on enumerated objects than other named objects. This suggests that participants had two different counting strategies for objects, one based on subitizing and the other to counting objects. To confirm this, we will collect more data and conduct a qualitative analysis.

References

- Aunio, P. & Niemivirta, M. (2010). Predicting children's mathematical performance in grade one by early numeracy. *Learning and Individual Differences*, 20(5), 427–435.
- Batchelor, S., Inglis, M. & Gilmore, C. (2015). Spontaneous focusing on numerosity and the arithmetic advantage. *Learning and Instruction*, 40, 79–88.
- Baumanns, L., Pitta-Pantazi, D., Demosthenous, E., Lilienthal, A. J., Christou, C. & Schindler, M. (2024). Pattern-Recognition Processes of First-Grade Students: An Explorative Eye-Tracking Study. *International Journal of Science and Mathematics Education*, 22, 1663–1682.
- Chatfield, M. & Mander, A. (2009). The Skillings-Mack test, *Stata J.*, 9(2), 299–305.
- Elliott, L., Silver, A. M., Imbeah, A. & Libertus, M. (2022). Actions may speak louder than words: Comparing methods of assessing children's spontaneous focusing on number. *Journal of Experimental Child Psychology*, 214.
- Gloor, N., Leuenberger, D. & Moser Opitz, E. (2021). Disentangling the Effects of SFON (Spontaneous Focusing on Numerosity) and Symbolic Number Skills on the Mathematical Achievement of First Graders. A Longitudinal Study. *Frontiers in Education*, 6.

- Hannula, M. & Lehtinen, E. (2005). Spontaneous focusing on numerosity and mathematical skills of young children. *Learning and Instruction*, 15(3), 237–256.
- Hannula, M. M., Lepola, J. & Lehtinen, E. (2010). Spontaneous focusing on numerosity as a domain-specific predictor of arithmetical skills. *Journal of Experimental Child Psychology*, 107(4), 394–406.
- Hannula-Sormunen, M., Lehtinen, E. & Räsänen, P. (2015). Preschool Children's Spontaneous Focusing on Numerosity, Subitizing, and Counting Skills as Predictors of Their Mathematical Performance Seven Years Later at School. *Mathematical Thinking and Learning*, 17(2–3), 155–177.
- Jordan, N. C., Kaplan, D., Ramineni, C. & Locuniak, M. N. (2009). Early Math Matters: Kindergarten Number Competence and Later Mathematics Outcomes. *Developmental Psychology*, 45(3), 850–867.
- Just, M. A. & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87(4), 329–354.
- Martin, E. (1974). Saccadic suppression: A review. *Psychological Bulletin*, 81, 899–917.
- Nanu, C., Laakkonen, E. & Hannula-Sormunen, M. (2020). The effect of first school years on mathematical skill profiles. *Frontline Learning Research*, 8(1), 56–75.
- Nanu, C. E., McMullen, J., Munck, P. & Hannula-Sormunen, M. M. (2018). Spontaneous focusing on numerosity in preschool as a predictor of mathematical skills and knowledge in the fifth grade. *Journal of Experimental Child Psychology*, 169, 42–58.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372–422.
- Schindler, M., Schovenberg, V. & Schabmann, A. (2020). Enumeration Processes of Children with Mathematical Difficulties: An Explorative Eye-Tracking Study on Subitizing, Groupitizing, Counting, and Pattern Recognition. *Learning Disabilities: A Contemporary Journal*, 18(2), 193–211.
- Sprenger, P. & Benz, C. (2020). Children's perception of structures when determining cardinality of sets – results of an eye-tracking study with 5-year-old children. *ZDM*, 52, 753–765.
- Trick, L & Pylyshyn, Z (1994). Why Are Small and Large Numbers Enumerated Differently? A Limited-Capacity Preattentive Stage in Vision. *Psychological Review*, 101(1), 80–102.
- Torbeyns, J., Bojorque, G., Van Hoof, J., Van Nijlen, D. & Verschaffel, L. (2017). Unique contribution of Ecuadorian kindergartners' spontaneous focusing on numerosity to their early numerical abilities. *British Journal of Developmental Psychology*, 36(2), 299–312.

HOW STUDENTS EVALUATE DIFFERENT TASK FORMATS? AN EXPLORATORY STUDY IN A LARGE MATHEMATICS COURSE FOR BUSINESS STUDENTS

Birte Reich¹, Michael Liebendörfer¹, Andreas Eichler², Mirko Schürmann¹, Nadja Stroh-Maraun¹ and Raphael Müller¹

¹Paderborn University, Germany, ²University of Kassel, Germany

Digital tasks are increasingly used in university mathematics but there is little evidence about whom they serve in particular. In a large business mathematics course, we compared student evaluations of (1) traditional small-group tasks, (2) digital STACK tests to be solved at home and (3) digital multiple-choice quizzes used live in tutorials. We explored students' use and evaluation of teaching offers in relation to their mathematical skills and motivation. We found that small-group tasks were preferred by frequent course participants and STACK tasks appealed to high-achieving students unlike those who favored quizzes. For both STACK tasks and the quizzes, the immediate feedback was considered particularly useful. We discuss developing and implementing diverse task formats to better cater students' diversity.

TASKS IN MATHEMATICS TEACHING

Mathematics tasks play an important role in effective teaching, serving as tools to help students connect with the material, practice skills and gain feedback. In university courses where students are free to choose how much time and effort to put into tasks, these tasks should be appealing and correspond to the learners' abilities and needs. In recent years, digital task formats have become widely used offering flexibility for homework and live use in lectures and tutorials. They can activate students and give both individual feedback to the learners and feedback on students' understanding to the lecturer. Even when used in large courses, the lecturer's workload remains limited. Given the growing diversity in student needs, especially with students working remotely, incorporating digital tasks might support a broader range of students.

However, there is currently a lack of understanding about why some students successfully use certain task formats while others struggle. While there are numerous positive reports on the use of digital tasks in small-group settings, little research has explored the implementation of different formats in large courses, despite being designed for such contexts. This study investigates students' preferences for three task formats in a large business mathematics lecture at Paderborn University: traditional paper-based tasks in small-group tutorials, open STACK tasks on an online platform, and live multiple-choice quizzes in large tutorials. We analyze students' participation, task usage, course evaluations, and their mathematical skills and motivation to identify preferences for specific formats.

LITERATURE REVIEW

Tasks are an important part of mathematics teaching, varying in content and implementation. Common formats include multiple-choice (MC) questions, which provide predefined answers, and constructed response (CR) items where the solutions must be elaborated by the students themselves. Their nature is slightly different, as MC allows reverse-solving strategies and offers short feedback and learning progress comparison directly whereas CR allows more detailed feedback but permits only one solving direction (Shepard, 2008).

Digital tasks provide immediate feedback, often through response distribution charts that help students assess their performance and learn from peers' mistakes (Núñez-Peña & Bono, 2022). Online formats, such as STACK in Moodle, offer mathematical problems, interpret submitted mathematical expressions, analyze students' answers and offer feedback (Sangwin, 2010). The feedback function in particular makes STACK valuable for formative assessments, as students can quickly correct their mistakes and make progress in their learning (Sangwin, 2015). Nevertheless, digital task formats have their own drawbacks, such as difficulties with input (Sangwin, 2015).

To the best of our knowledge, there is no specific literature on students' evaluation of different task formats in university mathematics. We therefore explore this issue with a focus on students' self-regulated learning (Boekaerts, 1996). In doing so, we are following the idea that the students themselves largely determine how much they use voluntary activities. The framework of self-regulated learning supposes that students are able to willingly control what they do. Their choices involve students' evaluation of learning opportunities, their own abilities, and their motivation. The latter may include unpleasant aspects such as mathematics anxiety (Götz, 2004).

The perspective of self-regulated learning led to interesting findings in university mathematics. In their qualitative study, Göller et al. (2024), demonstrated how important short-term goals (such as passing exams), long-term goals (e.g., relating to the student's future profession) and students' well-being (including negative forms like anxiety) can be for working on tasks during studies. Kempen and Liebendörfer (2022) found during the COVID-19 pandemic that students who preferred digital learning resources over traditional learning resources had a rather negative relationship to mathematics and assessed their own mathematical abilities worse. However, in contrast to our study, their preferred resources were not closely linked to the lecture.

THE PRESENT STUDY AND RESEARCH QUESTION

We analyzed a course in which three types of tasks (classical paper-based, digital STACK homework, live MC quizzes) were used. Our study took place in the winter semester 2023/24 in a first-semester mathematics course for business students (1072 assigned students) at a medium-sized German university. The course focused on procedural mathematics like calculus but not proof or problem solving. It consisted of a weekly lecture, a small-group tutorial and a full-class tutorial as well as homework.

Traditional paper-based tasks were solved and discussed in the small-class tutorial that emphasized understanding the tutor demonstrations. Participating students received the solutions for exam preparations.

STACK tests were implemented as homework. Seven tests comprising of 8–10 tasks were provided online for a two-week period so that each student had access to the tasks. Students had an infinite number of attempts, with the solutions and individual feedback only being displayed after the final submission. By passing a STACK test, students received one bonus point for the exam (up to five), enhancing the grade only if the exam was passed.

Quizzes were implemented in the full-class tutorial where questions were clarified, and tasks were solved. The quizzes should enable peer discussions. The answer distribution was then displayed for anonymous self-assessment and only then the solution was revealed. The quizzes used the web-based multiple-choice system PINGO, which supports mathematical notation (Paderborn University, 2011). This format should actively engage students, promoting participation.

Neither the lecture nor the tutorials nor the homework was mandatory, and in fact for years many students have participated in this course on a limited basis. The lecturer innovated and used new digital task designs. We thus evaluated this course and aimed at characterizing students who evaluate specific task formats positively. Our research question was: “How do students use and evaluate the different task formats (traditional, STACK, MC) in relation to their mathematical skills and motivation?”

METHODS

To answer our research question, we conducted an exploratory study and developed a questionnaire in consultation with the lecturer. Our online survey was completed by 251 students, but not everyone answered all the questions. For data protection reasons, we have not asked for any personal characteristics. The survey was presented during the lecture and additionally via instructional e-mail and could be voluntarily accessed for the final four weeks of the semester. The questionnaire consisted mostly of single-item questions, but also of four scales.

We assessed students’ evaluation of the learning opportunities in the course (lecture, full-class tutorial, small-group tutorial and tasks, STACK tasks and MC quizzes) by asking them to evaluate them using an answer scale from very bad [1] to very good [5] or to give reasons for non-usage. To measure students’ usability of MC quizzes, we developed a scale with four items that asked if the quizzes were helpful or annoying, if they should give the solution directly (which they did not) and if the comparison with peers was beneficial. To assess the usability of STACK tasks, we developed a four-item scale that evaluated how well students managed the input process and the processing process and the perceived helpfulness of the feedback and of the input interpretation.

We further asked students about their participation in each course measure on a 5-point scale from never [1] to very often or very intensively [5]. To assess students' workload, we asked for the required hours/week and to rate the workload from very low [1] to very high [5].

We assessed students' mathematical skills in terms of their final school grade in mathematics, with [1] being the best and [6] the worst according to the German school grades. We further asked for a self-assessment of mathematical competence in relation to the other course members, whether they considered themselves among the best 25%, the next highest 25%, the next lowest 25% or the weakest 25%.

To assess motivational aspects, we asked students if their goal for the course is only to pass and that the grade is of minor importance (using a 5-point agreement scale). We further determined the students' view of the usefulness of mathematics for their future profession using a scale with four items from Isaev and Eichler (2022) and assessed mathematics anxiety with a three-item-scale from Götz (2004).

We present descriptive statistics and the Pearson correlations to describe the evaluation and participation in the course measures, mathematical skills and motivation in relationship to the evaluation of small-group tasks, STACK and the quizzes. The data were treated as metric on a scale ranging from 1 to 5. All four scales had a high internal consistency (Cronbach's $\alpha > .71$). Missing values were pairwise excluded from the analyses. The data on participation of the measures is subject to a strong ceiling effect; it was primarily students who took part in the survey who also made heavy use of the teaching measures. We therefore primarily analyze the data on the evaluation of the three task types instead of the data on their use.

RESULTS

The participation in all task formats was very high (Table 1), which seemed to result in a ceiling effect that might suppress correlations. Only the participation in small-group tasks was lower, with students citing time constraints or scheduling conflicts as reasons for non-participation. We thus focus on students' evaluation of the tasks formats. Students who evaluated one type of the tasks positively often evaluated others favorably as well and rated the teaching concept of the course better.

The better students evaluated the tasks, the more likely they were to perceive mathematics useful for them in their future career, with the strongest correlation observed for the traditional small-group tutorial (.32). All students who positively evaluated a task format experienced less anxiety about mathematics, with the negative correlation being most pronounced for STACK tasks (-.38) and least for quizzes (-.23).

The traditional paper-based tasks in the small-group tutorials were highly valued by students who consistently attended in-person course components, such as lectures and tutorials. The positive evaluation of these tasks had a strong correlation with the overall course evaluation (.47) and low mathematics anxiety (-.31). The correlation with mathematics skills was lower.

					Pearson Correlation with evaluation of		
	Item/scale [no. of items]	n	Mean	SD	Small- group tasks	STACK tasks	MC quizzes
Evaluation	Lecture	250	4.59	0.76	.19**	.02	.04
	Full-class tutorial	251	4.37	0.95	.08	.25	.04
	Small-group tasks	231	3.91	0.99	1	.35**	.23**
	STACK tasks	235	3.96	0.99	.35**	1	.38**
	STACK usability [4]	235	3.79	0.77	.46**	.59**	.37**
	MC quizzes	234	3.64	0.93	.23**	.38**	1
	MC quiz usability [4]	244	3.44	0.97	.18**	.12	.49**
	Course	246	3.72	0.93	.47**	.35**	.34**
Math. Participation	Lecture	250	4.59	0.76	.19**	.02	.04
	Full-class tutorial	251	4.37	0.95	.12	.08	.13*
	Small-group tutorial	251	3.96	1.22	.27**	-.04	-.06
	STACK tasks	251	4.37	1.11	.11	.36**	.03
	MC quizzes	251	4.07	1.16	.11	.07	.41**
	Workload (hours/week)	251	6.99	4.72	.10	.22**	.15*
	Workload evaluation	247	3.40	0.92	-.09	-.24**	-.16*
	Final school grade	212	2.60	0.89	-.18*	-.30**	-.12
Math. skills	Self-assessment	246	2.76	0.83	.15*	.32**	.33**
Motivation	The goal is to pass (grade not important)	248	3.61	1.38	-.16*	-.34**	-.20**
	Usefulness of mathematics [4]	248	2.98	0.86	.32**	.24**	.23**
	Math anxiety [3]	247	3.23	1.23	-.31**	-.38**	-.23**

* $p < .05$, ** $p < .01$ (2-tailed).

Table 1: Descriptive statistics and correlations of the items/scales

STACK tasks were evaluated positively when their usability was evaluated high. This meant students rated the difficulty appropriate, the input and processing process straightforward, and the feedback and interpretation aid for the input helpful according to the STACK usability scale. There was a correlation between positive evaluations of the STACK tasks and much time spent per week (.22) but the workload being

considered low (-.24). STACK tasks were primarily well rated by higher achievers in school (-.30), who also now evaluated their own mathematics skills better (.32). Students with less mathematics anxiety also provided better evaluations (-.38).

MC quizzes were particularly popular for enabling students to assess their learning progress relative to peers. Reportedly, the fact that the solution was not displayed immediately had no significant influence on the rating. For the students who evaluated MC quizzes positively we could see a slight correlation to a low workload (-.16) and high amount of time they were using per week (.15). They rated their own mathematics skills higher (.33), although no correlation to the school grades was observed. Their goal was rather to achieve a good grade than just to pass (-.20). The more students had taken part in the quizzes, the more they evaluated them positively (.41).

DISCUSSION

Our aim was to explore how students use and evaluate different task formats in university mathematics in relation to their skills and motivation. We investigated a course in which three task types were implemented: traditional paper-based tasks that were completed in small groups, online STACK tasks with the bonus point system and multiple-choice (MC) quizzes. We collected data using a questionnaire that captured students' participation, evaluation of the measures, mathematics skills and motivation, which we analyzed quantitatively.

In general, students who considered mathematics as useful evaluated all three task formats more positively, confirming that one's strategies are chosen according to one's beliefs (Boekaerts, 1996). They were also more motivated to achieve high grades and had less mathematics anxiety (Götz, 2004). Furthermore, students who evaluated the task formats positively also assessed their mathematical skills as rather good. Contrary to expectations (e.g. Kempen & Liebendörfer, 2022), digital task formats, such as STACK tasks and MC quizzes, seem to appeal to roughly the same students who prefer classical tasks in a small-group tutorial. This suggests that we may still lose students-at-risk with digital tasks and contradicts the assumption that digital formats are more accessible because you get feedback without embarrassing yourself in front of others.

Traditional paper-based small-group tasks were well-rated by students who also participated in other attendance options, such as the lecture or full-class tutorial. This indicates a preference for face-to-face learning or a commitment to attending all sessions. As this tutorial was held at certain times, some students did not take part in the tutorial due to overlaps or no time. Compared to digital formats there are different correlations with mathematical self-assessment, workload and their learning goals suggesting that other students are addressed through this task format.

STACK tasks were evaluated positively by high-achieving students with strong mathematics skills. Students appreciated their feedback and complexity aligning with the findings of Sangwin (2015). As these were CR tasks, this supports the findings of Shepard (2008). The students who rated the STACK tasks well also stated that their

workload was higher, which can be explained by the fact that these tasks had to be completed from home in addition to the attendance dates and therefore did not take place within the specified course times. The high participation in STACK tasks can be related to the external motivation factor, as bonus points for the final exam were achieved with these tasks. As these were only counted if the exam was passed anyway, it explains the high correlation between positive evaluations of the STACK tasks and the goal of achieving good grades.

MC quizzes appealed broadly due to their progress comparison feature. Students who evaluated MC quizzes positively rated their performance better than their fellow students but did not necessarily have better school grades. This suggests that MC quizzes may appeal equally to students of varying mathematical abilities, but the opportunity to compare their results with peers may contribute to their appeal for high-performing students. Interestingly, not all students in the full-class tutorial, where the quizzes are integrated, rated quizzes favorably, whereas the quizzes participants did. This indicates that some participants did not take part in the quizzes, potentially due to a lack of alignment with their preferences for task formats. Anyway, we could not reach everyone in the room with digital tasks and in future, we should seek ways to do so.

These findings underscore the importance of offering diverse task formats to address the heterogeneous needs of students in large mathematics courses and highlights the need for future efforts to develop strategies that ensure broader inclusivity and participation.

Limitations and implications for research

In this study, the items were only roughly selected to obtain an overview of potential differences in utilization and evaluation regarding the task formats. Future research should refine the questionnaire to include detailed items on students' motivations and thought processes for better precision.

Key findings, such as the preference for STACK tasks among high-performing students, raise questions about whether this preference stems from their mathematical abilities or specific task features like delayed feedback. Similarly, the discrepancy between full-class tutorial participants and quiz users highlights the need to explore students' perceptions of task formats and how this affect engagement. Another limitation is the ceiling effect in participation: the scale mean values are so high that we were probably unable to capture the true variance between the students well in the upper range. This probably means that some correlations could not be statistically visible.

ACKNOWLEDGMENT

We would like to thank the Stiftung Innovation in der Hochschullehre for funding our LLV.HD project, of which this study is a part.

References

- Boekaerts, M. (1996). Self-regulated learning at the junction of cognition and motivation. *European Psychologist*, 1(2), 100–112. <https://doi.org/10.1027/1016-9040.1.2.100>
- Göller, R., Gildehaus, L., & Lahdenperä, J. (2024). Students' self-regulated learning of university mathematics in different learning environments. *International Journal of Mathematical Education in Science and Technology*, 1–24. <https://doi.org/10.1080/0020739X.2024.2341035>
- Götz, T. (2004). *Emotionales Erleben und selbstreguliertes Lernen bei Schülern im Fach Mathematik* [Emotional experience and self-regulated learning in mathematics students]. München: Utz.
- Isaev, V., & Eichler, A. (2022). Der Fragebogen zur doppelten Diskontinuität. In S. Halverscheid, I. Kersten & B. Schmidt-Thieme (Hrsg.), *Bedarfsgerechte fachmathematische Lehramtsausbildung. Analyse, Zielsetzungen und Konzepte unter heterogenen Voraussetzungen* [Needs-orientated mathematics teacher training. Analysis, objectives and concepts under heterogeneous conditions] (S. 321–338). Springer. https://doi.org/10.1007/978-3-658-34067-4_18
- Kempen, L., & Liebendörfer, M. (2021). University students' fully digital study of mathematics: an identification of student-groups via their resources usage and a characterization by personal and affective characteristics. *Teaching Mathematics and its Applications: An International Journal of the IMA*, 40(4), 436–454. <https://doi.org/10.1093/teamat/hrab020>
- Núñez-Peña, M.I., & Bono, R. (2022). Using error analysis sheets to improve learning in higher education. *8th International Conference on Higher Education Advances (HEAd'22)*. <https://doi.org/10.4995/head22.2022.14656>
- Paderborn University (2011). *PINGO*. <https://www.uni-paderborn.de/lehre/lehrinnovationen/lehrprojekte/pingo>
- Sangwin, C. J. (2010). *Who uses STACK? A report on the use of the STACK CAA system* (Tech. Rep). The Mats, Stats and OR Network, School of Mathematics, The university of Birmingham.
- Sangwin, C. (2015). Computer Aided Assessment of Mathematics Using STACK. In: Cho, S. (eds) *Selected Regular Lectures from the 12th International Congress on Mathematical Education*. Springer, Cham. https://doi.org/10.1007/978-3-319-17187-6_39
- Shepard, L. A. (2008). Commentary on the National Mathematics Advisory Panel Recommendations on Assessment. *Educational Researcher*, 37(9), 602–609. <https://doi.org/10.3102/0013189X08328001>

UNDERSTANDING THE SOLUTION SET OF SYSTEMS OF LINEAR EQUATIONS THROUGH GEOMETRIC VISUALIZATION

Elizabeth Rozas-Torres^{1,2}, Andrea Cárcamo², and Josep Maria Fortuny¹

¹ Universtitat Autònoma de Barcelona, ²Universidad Austral de Chile

The objective of this study is to determine how the geometric visualization of a system of linear equations with three variables facilitates the understanding of its solution sets among university engineering students. Based on visualization processes and modes of thinking in Linear Algebra, design-based research was conducted with 25 first-year university students enrolled in a Linear Algebra course. The participants completed two learning tasks on a system of linear equations involving geometric visualization. The results indicate that geometric visualization, complemented by the parametric representation of a line's equation, enables students to understand the nature and structure of the solution sets of SLE in \mathbb{R}^3 .

INTRODUCTION

Visualization is recognized as an important element of mathematical reasoning (Presmeg, 2020). Several researchers highlight the important role of visual resources in the construction of mathematical concepts, as they help to develop an intuition for the abstract (Arcavi, 2003; Yilmaz & Argun, 2018). In Linear Algebra, geometric interpretations were fundamental to its historical development and enabled a better understanding of its methods (Winter, 1992). Nevertheless, current teaching practices often minimize this geometric dimension when introducing new concepts.

Research in Linear Algebra demonstrate that university students often struggle to solve and conceptualize the solution set of a system of linear equations (SLE) using matrix methods, especially in cases involving infinite or empty solutions (Ochoviet, 2009; Rodríguez et al., 2019). This problem is attributed to a mechanical and abstract teaching approach (Rodríguez et al., 2022), that is disconnected from real-life or familiar situations for the student. Considering the aforementioned, we designed two learning tasks that integrate: the geometric visualization of linear equations with three variables, the relative positions between planes, and the parametric equation of a line. Therefore, the objective of our research is to determine how geometric visualization facilitates the understanding of the solution sets of SLEs with three variables among university engineering students.

THEORETICAL FRAMEWORK

Visualization processes

Visualization, according to Presmeg (2020, p. 900), comprises:

“...the ability, the process and the product of creation, interpretation, use of and reflection upon pictures, images, diagrams, in our minds, on paper or with technological tools, with

the purpose of depicting and communicating information, thinking about and developing previously unknown ideas and advancing understandings”.

Visualization involves two processes (Bishop, 1989): the visual processing of information (VPI), which converts external information into mental images, and the interpretation of figurative information (IFI), which analyzes these images to extract relevant information. Both processes play a crucial role in regulating and managing the connection between the external information perceived by the student and the internal information generated through their mental activity.

Modes of thinking

According to Sierpinska (2000), three modes of thought coexist in Linear Algebra: synthetic-geometric, analytical-arithmetic and analytical-structural. Each of these modes uses a specific system of representations. The synthetic-geometric mode corresponds to geometric objects as they are perceived by our senses. For example, one is in the synthetic-geometric mode if one thinks about the possible solutions for the SLE by visualizing the possible relative positions of three planes in space. In the analytical-arithmetic mode, the equations that form a SLE can be solved by using different methods. To interpret the symbols as equations it is necessary some background, for example, if one thinks in terms of the possible results of a row reduction of an augmented matrix. In the analytical-structural mode, the emphasis is on the properties of the objects, for example, thinking about the conditions in which a SLE has an unique, empty or infinite solution. Sierpinska (2000) emphasizes that the ability to switch between different representations is essential for deep learning in Linear Algebra. In particular, geometric visualization stands out as a powerful tool to facilitate the understanding of abstract concepts and their connections.

METHODOLOGY

The methodology employed was a Design-Based Research consisting of three phases: design, experimentation, and retrospective analysis (Bakker, 2018). In the design phase, we developed a hypothetical learning trajectory (Simon, 2020; Cárcamo et al. 2021) with 4 learning tasks. Learning tasks 1 and 2 incorporated geometric visualization: the first addressed 2×3 SLE, activating prior knowledge about linear equations and relative positions between planes, while the second explored 3×3 SLE through different arrangements of three planes. Learning tasks 3 and 4 extended the work to more general SLE without a visual component.

In the experimentation phase, we conducted two cycles where we applied the hypothetical learning trajectory in a Linear Algebra course. In this study, we focused on the second experiment cycle, in which 25 first-year engineering students enrolled in Linear Algebra at a Chilean university participated. According to the Chilean high school curriculum, they had studied 2×2 SLE, solving them by substitution, reduction, and graphing methods. Additionally, they had studied plane and line equations in space in their geometry course, which is a prerequisite for taking the Linear Algebra course.

The experiment was conducted in 5 sessions (60 to 90 minutes each), collecting written protocols from students' learning tasks, audio recordings, and notes. In the analysis phase, we transcribed the audio recordings of discussions among students and the written protocols. We performed an inductive qualitative analysis, identifying emerging categories. For each category, representative episodes that evidenced students' reasoning were selected, triangulating data from dialogue transcriptions and their written protocols. The validity of the analysis was ensured through iterative review among researchers until a consensus was reached on the interpretation of it. In this research, we present the results of the data analysis of the first two learning tasks of the hypothetical learning trajectory due to their emphasis on geometric visualization.

RESULTS

The results highlight three key aspects in understanding the solution set of SLE through geometric visualization: (1) activation of prior knowledge of the linear equation's geometric representation in \mathbb{R}^3 , (2) the articulation between geometric and arithmetic representations of the solution set, and (3) the use of the parametric equation of the line to express the solution set. Below, we describe and illustrate these aspects using students S1 and S2 as examples.

Activation of previous knowledge on the geometric representation of the linear equation

Visualizing a linear equation as a three-dimensional plane

During learning task 1, students activated their prior knowledge of linear equation with three variables at the time of exploring equation $Ec1: x + 2y - z = 3$ and identified its geometric representation as well as its number of solutions. The following dialogue between S1 and S2 shows that the students are in the synthetic-geometric mode to justify the infinite solutions of the linear equation, as their justifications refer to the infinite points that belong to the plane:

- 1 S1: (referring to $Ec1: x + 2y - z = 3$) represents a plane and has infinite solutions.
- 2 S2: Why does it have infinite solutions?
- 3 S1: Because the plane is infinite, I suppose.
- 4 S2: Because it has infinite points.
- 5 S1: Supposedly this stretches out [referring to the graph of the plane].

This dialogue provides evidence of how students connected the geometric properties of the plane to the cardinality of the linear equation solution set, highlighting the role of visualization in understanding the solution. Instead of resorting to algebraic arguments (the infinite points that satisfy the equation), the students based their reasoning on the geometric interpretation of the plane as a place of infinite points.

Intersection of two planes and the parametric equation of a line

When the students graphed the given equations $Ec1: x + 2y - z = 3$ and $Ec2: 2x + 3y + z = 1$ by using GeoGebra, they observed that these equations represented two intersecting planes. Regarding this, S1 indicated that *"both graphs intersect in a line"*. Subsequently, the students activated their prior knowledge of the line as a geometric locus and how to express its parametric equation by algebraically solving a 2×3 SLE using the reduction or substitution method. For example, student S1 combined both methods to solve the variables y, z based on x . Then, S1 assigned to x the value of a parameter t in \mathbb{R} , and obtained the following expressions: $y = \frac{4}{5} - \frac{3}{5}t$ and $z = -\frac{7}{5} - \frac{1}{5}t$. Thus, S1 was able to express the parametric equation of the line resulting from the intersection of the given planes (Figure 1).

$$\begin{aligned} & \begin{cases} x + 2y - z = 3 \\ 2x + 3y + z = 1 \end{cases} \\ & \begin{aligned} & x + 2\left(\frac{4}{5} - \frac{3}{5}x\right) - z = 3 \\ & x + \frac{8}{5} - \frac{6x}{5} - z = 3 \\ & -\frac{4x}{5} - z = \frac{7}{5} \\ & -z = \frac{7}{5} + \frac{4}{5}x \\ & z = -\frac{7}{5} - \frac{4}{5}x \end{aligned} \\ & \begin{aligned} & 3x + 5y = 4 \\ & 5y = 4 - 3x \\ & y = \frac{4}{5} - \frac{3}{5}x \end{aligned} \\ & \begin{cases} x = t \\ y = \frac{4}{5} - \frac{3}{5}t \\ z = -\frac{7}{5} - \frac{1}{5}t \end{cases} \quad t \in \mathbb{R} \end{aligned}$$

Figure 1: Student S1's written protocol in determining the parametric equation of the line resulting from the intersection of the given planes.

The students, after determined the parametric equation of the line l , represented the set of all points belonging to the intersection of the given planes as a subset of \mathbb{R}^3 and two different representations predominated. For example, student S1 wrote $\{(x, y, z) \in \mathbb{R}^3 \mid x = t, y = \frac{4}{5} - \frac{3}{5}t, z = -\frac{7}{5} - \frac{1}{5}t, t \in \mathbb{R}\}$, meanwhile their classmates wrote $\left\{\left(\frac{4}{3} - \frac{5t}{3}, t, \frac{t}{3} - \frac{5}{3}\right), t \in \mathbb{R}\right\}$.

During this process, students were able to use their prior knowledge regarding the parametric equation of the line and its set notation, showing their ability to transition between geometric and analytical-arithmetic modes.

Empty intersection

On the other hand, by analysing equations $Ec1: x + 2y - z = 3$ y $Ec3: -3x - 6y + 3z = 10$, student S1 expressed to their classmate that "it is no necessary to graph, because if we amplify by 3 here [and its added] (showing equation $Ec1$) it becomes zero, so zero solutions". Student S1 algebraically deduced the nonexistence of intersection points by observing that one equation was a multiple of the other one. Consequently, student S1 graphed the equations and could verify what had already been identified. S1 could conclude that *"they are planes parallel to each other, therefore they do not have common points"*.

Articulation between the geometric and arithmetic representation of the solution set

Students established connections between the geometric and algebraic of the SLE when they linked their previous work starting from analytic geometry (intersections of planes) with the arithmetic solutions of the SLE formed by the 2 linear equations. For example, student S1 identified that the equation involved in each SLE were those previously worked on ($Ec1$, $Ec2$ and $Ec1$, $Ec3$). Then, S1 sketched each SLE geometrically (Figure 2, second column).

SEL de 2×3	Representación geométrica	Cantidad de soluciones	Conjunto solución
$\begin{cases} x+2y-z = 3 \\ 2x+3y+z = 1 \end{cases}$		infinitas soluciones	$\{(x,y,z) \in \mathbb{R}^3 / x=2, y=\frac{2}{3}-\frac{2z}{3}, z \in \mathbb{R}\}$
$\begin{cases} x+2y-z = 3 \\ -3x-6y+3z = 10 \end{cases}$		0 soluciones	$C_s = \emptyset$

Figure 2: Student S1's written protocol from learning Task 1 on the geometric representation, number of solutions, and solution set of two 2×3 SLEs.

As it is demonstrated in the first 2×3 SLE case (Figure 2, first row), student S1 drew two intersecting planes, indicating that it had infinite solutions. S1 stated: *"Taking this into account, what we obtained here [indicating what it is in Figure 1] ... is the line."*. Student S1 realized that the set of all points that are in the previously obtained line, by intersecting planes from equations $Ec1$ and $Ec2$, is in fact the solution set of the first suggested SLE and wrote it down as a solution set of the first SLE (Figure 2).

In the second 2×3 SEL case (Figure 2, second row), S1 drew two parallel planes and, although with some inaccuracies with the empty solution set notation, which are due to difficulties with set notation, student S1 identified and expressed that the second SEL has zero solutions. Additionally, S1 provided evidence of understanding the relation between geometric representation and algebraic solution result when stated: *"If the graphic representation creates an intersection, it will have infinite solutions. That is because those solutions are given by the line. If the graphic representation is parallel, it will not have solutions because it does not generate an intersection"*. Student S1 identified that the relative positions between two planes correspond to a 2×3 SLE, which can have infinite solutions or an empty solution. We deduce that the student managed to transition from the geometric mode to the structural analytical mode by relating the different representation registers of the SLE and its solutions. In other words, the student understand that happens with the SLE when visualizing it geometrically; that means, if the two planes intersect, SLE has infinite solutions. Otherwise, if the planes are parallel, SLE will not have solutions.

For learning task 2, the geometric, matrix, and analytical representation mode of a 3×3 SLE were addressed. By visualizing the relative positions between 3 planes (Figure

3), students inferred the 3 types of solutions that a 3x3 SLE can have (unique, infinite, empty).

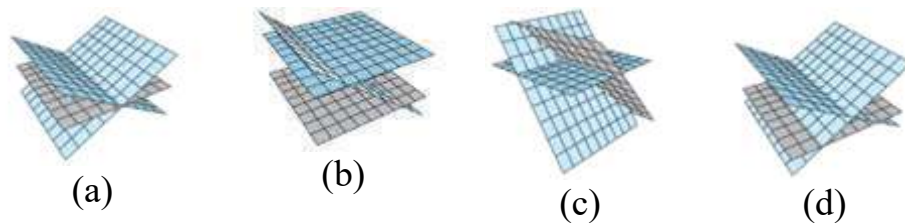


Figure 3: Examples of relative positions between the planes presented in learning task 2.

For the case presented in Figure 3(a), students mentioned it has infinite solutions because the three planes intersect in a single line. Next, in Figure 3(c) they indicated it has a unique solution because the planes intersect in a single point. Meanwhile, for cases presented in Figures 3(b) and 3(d) they mentioned that they do not have any solution because none of the three planes intersect in a single line or point.

Using the parametric equation of a line to write the solution set of a SEL

During learning task 2, students solved a SLE matrix form using their prior knowledge related to the parametric equation of the line to express the solution set as a subset of \mathbb{R}^3 . For example, S1 (Figure 4) wrote the given SLE in augmented matrix form, solved the SLE using basic row operations and then obtained a SLE equivalent to the initial SLE. Then, using the substitution method, S1 solved the variables x , z based on

$$\begin{pmatrix} 1 & -1 & 0 & -1 \\ -3 & 2 & 4 & 0 \\ 6 & -5 & -4 & -3 \end{pmatrix} \xrightarrow[t_2 + 3t_1]{t_3 - 6t_1} \begin{pmatrix} 1 & -1 & 0 & -1 \\ 0 & -1 & 4 & -3 \\ 0 & 1 & -4 & 3 \end{pmatrix} \xrightarrow{t_2 \leftrightarrow t_3} \begin{pmatrix} 1 & -1 & 0 & -1 \\ 0 & 1 & -4 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\begin{cases} x - y = -1 \\ -y + 4z = -3 \\ 0x + 0y + 0z = 0 \end{cases} \quad \begin{cases} x = -1 + y \\ -y + 4z = -3 \\ y = 4 - x \end{cases} \quad \begin{cases} x = -1 + t \\ y = t \\ z = -\frac{3}{4} + \frac{1}{4}t \end{cases}, t \in \mathbb{R}$$

$$S = \left\{ (-1+t, t, -\frac{3}{4} + \frac{1}{4}t) \mid t \in \mathbb{R} \right\}$$

variable y . Following, student S1: assigned the parameter t to variable y , created a parametric equation of the line, and wrote the solution set in terms of this parameter. In this process, that S1 carried out to express the solution, the student relied on the geometric visualization developed in learning task 1, specifically the parametric equation of the line, although without the need for graphing.

Figure 4: Student S1's written protocol while solving an SLE using matrix methods.

CONCLUSIONS AND DISCUSSION

Our research determined how the geometric visualization of an SLE with three variables facilitates the understanding of its solution sets among university students. The results show that the visualization processes (Bishop, 1989) were activated sequentially. First, the VPI enabled students to transform graphical representations of planes into mental images of how they intersect (or do not) in three-dimensional space. Then, the IFI worked on these representations to generate a meaning. Based on this visualization, students analyzed the relationships between planes (*Do they form a line of intersection? Do they have no common points?*) and connected this information with possible algebraic solutions.

The geometric visualization of SLEs with three variables proved to be an effective tool for facilitating the understanding of SLE solutions. Students were able to establish connections between visualization using software and their mental representations, allowing them to identify and characterize solution sets algebraically based on geometric intersections. This ability to transition between geometric and algebraic representations confirms Sierpinska (2000) claim that geometric visualization stands out as a powerful tool for facilitating the understanding of abstract concepts and their connections. Furthermore, geometric visualization emerges as an inclusive teaching practice that supports the construction of SLE solutions, as it allowed students to progressively develop their understanding from concrete spatial interpretations to abstract algebraic formulations.

Additional information

This work is presented in the context of the doctoral program in education at the Universitat Autònoma de Barcelona.

This research is part of the FONDECYT regular research project N°1241155 supported by ANID of Chile.

References

- Arcavi, A. (2003). The role of visual representations in the learning of mathematics. *Educational studies in mathematics*, 52(3), 215-241.
- Bakker, A. (2018). *Design research in education: A practical guide for early career researchers*. Routledge.
- Bishop, A. J. (1989). Review of research on visualization in mathematics education. *Focus on learning problems in mathematics*, 11(1), 7-16.
- Cárcamo, A., Fortuny, J. M. y Fuentealba, C. (2021). Las trayectorias hipotéticas de aprendizaje: un ejemplo en un curso de álgebra lineal. *Enseñanza de las Ciencias*, 39(1), 45-63. <https://doi.org/10.5565/rev/ensciencias.2857>
- Ochoviet, C. (2009). *Sobre el concepto de solución de un sistema de ecuaciones lineales con dos incógnitas* (Tesis doctoral, Centro de Investigación en Ciencia Aplicada y Tecnología

Avanzada, Unidad Legaria). Montevideo. Repositorio digital IPN
<http://repositoriodigital.ipn.mx/handle/123456789/11405>

- Presmeg, N. (2020). Visualization and Learning in Mathematics Education. In S. Lerman (ed.) *Encyclopedia of Mathematics Education*. Springer
- Rodríguez Jara, M. A., Mena Lorca, A., Mena-Lorca, J. J. F., Vásquez Saldías, P., & Del Valle Leo, M. E. (2019). Construcción cognitiva del conjunto solución de un sistema de ecuaciones lineales con dos incógnitas. *Enseñanza de las ciencias*, 37(1), 0071-92.
- Rodríguez, M., Mena-Lorca, A., Gregori, P., Vásquez, P., del Valle, M., & Parraguez, M. (2022). Comprensión del conjunto solución de un sistema de ecuaciones lineales de dos incógnitas: un estudio de casos. *Educación matemática*, 34(3), 163-193.
- Simon, M. (2020) Hypothetical learning trajectories in mathematics education. En: Lerman, S. (ed.). *Encyclopedia of Mathematics Education* (pp. 354-357). Springer.
- Sierpinska, A. (2000). On some aspects of students' thinking in linear algebra. In J.-L. Dorier (Ed.), *On the teaching of linear algebra* (pp. 209-246). Springer, Dordrecht.
- Winter, D. J. (1992). *Matrix Algebra*. Macmillan Publishing Company, U.S.A.
- Yilmaz, R. & Argun, Z. (2018). Role of visualization in mathematical abstraction: The case of congruence concept. *International Journal of Education in Mathematics, Science and Technology (IJEMST)*, 6(1), 41-57. DOI:10.18404/ijemst.328337

This contribution has been withdrawn from the conference

This contribution has been withdrawn from the conference

This contribution has been withdrawn from the conference

This contribution has been withdrawn from the conference

This contribution has been withdrawn from the conference

This contribution has been withdrawn from the conference

This contribution has been withdrawn from the conference

This contribution has been withdrawn from the conference

COMPONENTS OF COMPUTATIONAL THINKING DEMONSTRATED BY PRESERVICE MATHEMATICS TEACHERS DURING A MATH MODELLING TASK

Patricio Santibáñez, Bárbara Bustos-Osorio, Josefa Castillo-Funes

Pontificia Universidad Católica de Valparaíso

Currently, the development of mathematical competence is fundamental, and for that reason, Mathematical Modeling and Computational Thinking play an important role, as both allow solving problems in the real world. This research adopts a qualitative approach and describes the components of CT as demonstrated by preservice secondary mathematics teachers during an MM task, during a clinical interview. Findings reveal that components such as decomposition, abstraction, algorithm design, and debugging naturally emerge during the task-solving process, while iteration and generalization require more structured pedagogical interventions.

INTRODUCTION

In a world shaped by technology, where data and multidisciplinary work grow, integrating Mathematical Modelling (MM) and Computational Thinking (CT) is key in mathematics education (Leung et al., 2021). Niss and Blum (2020) highlight that the mathematics enable to address questions from extra-mathematical contexts, making the development of mathematical competence crucial. MM is one of these competences, involving the identification of real-world problems, translating them into mathematical terms, and validating solutions (Greefrath et al., 2022). On the other hand, CT is necessary in 21st-century education, enabling learners to solve real-world problems from new perspectives (Bocconi et al., 2022; Bravo-Preciado et al., 2024). Wing (2006) asserts CT is as fundamental as reading, writing or arithmetic, an idea increasingly reflected in international curricula (Abelson & Kong, 2024).

From this perspective, the connection between MM and CT is evident, as both equip students to address real-world challenges. Lehman (2024) suggests MM fosters algorithm development, a core CT component (Shute et al., 2017). These algorithms serve as formal models extending beyond the initial problem context.

However, limited research examines students' responses to MM tasks involving algorithm design, limiting teachers' ability to identify algorithmic thinking (Lehman, 2024). Ang (2021) notes that while ICT's role in MM has been widely studied, CT's role remains understudied. Exploring this relationship could enhance task design and skill development. In this context, this study aims to describe the components of CT as demonstrated by preservice secondary mathematics teachers during an MM task.

THEORETICAL/CONCEPTUAL FRAMEWORK

The study is framed in the theoretical constructs of MM and CT as described below.

Mathematical modelling

MM is described as a cyclical process consisting of various phases (Kaiser et al., 2023). The literature highlights different MM cycles, noting that "depending on which aspects of the key stages in the basic modeling cycle are of particular interest in each context, the cycle can be expanded by elaborating on these aspects" (Niss & Blum, 2020, p. 14). Schukajlow et al. (2023) add that "the notion of a 'cycle' emphasizes the cyclical nature of modeling" (p. 260), as resolving an MM situation often requires modifying the approach or model according to the proposal.

This study considers the cycle proposed by Stillman et al. (2007) (Figure 1), which emphasizes transitions between phases and iterative processes depending on the modeler's needs when addressing real-world situations.

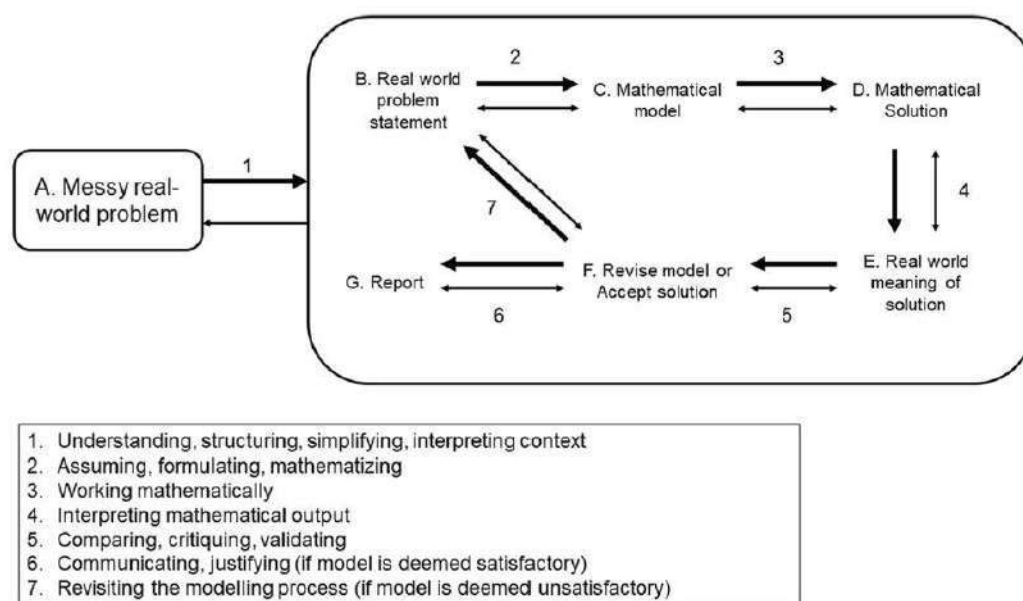


Figure 1 Mathematical Modelling Cycle (Stillman et al., 2007).

Computational Thinking

Regarding to CT, it is important to note that no consensus exists on its definition (Peracaula-Bosch & González-Martínez, 2024). However, Wing, in her attempts to define it, has proposed several revisions to her initial definition. In particular, in 2011, she stated that "CT is the thought process involved in formulating problems and their solutions so that the solutions are represented in a form that can be effectively executed by an information-processing agent" (p. 20). Later, in 2014, she clarified that these agents can be human or machine. Since then, Wing has not updated her definition, but other researchers have characterized CT, particularly in the context of STEM education (Weintrop et al., 2016; Shute et al., 2017).

Weintrop et al. (2016) propose a taxonomy of CT-related activities in STEM courses, divided into four categories: Data Practices, Modelling and Simulation, Computational Problem Solving, and Systems Thinking. The modelling and simulation group, in particular, offers a promising approach for incorporating CT into mathematics

education. Furthermore, Shute et al. (2017) suggest that CT can be characterized through its components (Table 1).

Component	Definition
Decomposition	Breaking down a problem into manageable parts.
Abstraction	Data collection and analysis, pattern recognition, and modelling.
Algorithm	Designing logical and ordered instructions to solve a problem. Subcategories: algorithm design, parallelism, efficiency, and automation.
Debugging	Detecting and identifying errors and fixing them when a solution does not work as expected.
Iteration	Repeating design processes to refine solutions until the ideal result is achieved.
Generalization	Transferring CT skills to a wide range of situations/domains to solve problems effectively and efficiently.

Table 1 Components of CT. Based on the work of Shute et al. (2017)

METHODOLOGY

This research is guided by a qualitative perspective, employing an intrinsic case study approach (Stake, 2015). A convenience sampling method was used. The study participants included three preservice secondary mathematics teachers in their final year of training at a Chilean university. Two of them had taken an elective course on CT, but none had formal training in Discrete Mathematics.

Data was collected via a clinical interview lasting one hour, where participants collaboratively tackled an MM task (Figure 2), designed specifically for this study. In addition to this information, they were given a table listing 52 people, each associated with the person who presumably infected them, identified by codes (id00NN, with NN ranging from 01 to 52). Follow-up questions provided further insights.

Both the responses to the MM task and the follow-up questions were analysed using content analysis (Bardin, 2002) and based on the categories defined by the CT components proposed by Shute et al. (2017) (Table 1).

In Talca, a new variant of a highly contagious disease has been discovered, causing green-colored rashes, nausea, fever, headaches, and extreme fatigue. So far, around 1,000 cases have been detected. The police, the Ministry of Health, and the WHO are desperately working to locate patient zero, the individual who was the first infected and spread this serious disease. You are responsible for helping to identify this person (or these people) so that the Ministry of Health can isolate them and assist in efforts to develop containment and treatment protocols for the disease.

The Ministry of Health currently has some data based on information gathered from interviews with those who tested positive after a series of tests. The following table presents data on 52 individuals, showing with some certainty who infected whom.

Based on the data:

1. How could the way the virus spread be characterized?
2. How could it be determined who was (or were) the first to contract the disease?

Consider the following for your response:

- Be as clear and precise as possible in your answers.
- It is a good idea to use examples to clarify what you are trying to convey.
- If your response involves a procedure or algorithm, be precise in each step that must be followed to reach the solution.

Figure 2 MM Task (translated to English).

RESULTS

Below is a summary of the participants' actions, categorized by the CT components reflected in their work. The participants are referred to as Alejandra, Marcela, and Manuel to ensure anonymity while respecting their gender.

Decomposition

Using the definition of patient zero, "the first individuals infected who spread this disease", the participants identified two key properties: "being the first to be infected" and "spreading the disease". This allowed them to identify individuals 25 and 28 (for simplicity, the "id00" part of the identifier will be omitted) as potential patient zeros, since there was no information about who infected them, and they confirmed these individuals had infected others.

They also identified pairs, such as 1-8 and 21-19, who had infected each other and others but lacked information on who infected them, classifying them as a different type of patient zero. This led to discussions about data inconsistencies, theorizing that some individuals might not belong to the city of Talca or that data might come from a survey, introducing uncertainty. Pair 18-9, however, was excluded as they did not infect others, failing to meet both conditions.

Decomposition was evident in their approach to constructing an infection graph. They first analyzed individuals with no recorded infection source, then those with mutual infections, using this process to structure their solution.

Abstraction

According to Shute et al. (2017), abstraction involves data collection, pattern recognition, and modeling, aiming to extract the essence of a system. This was evident in the participants' discussion about patient zero characteristics. Alejandra centralized

the extracted information, drawing a directed graph ("tree diagram", Figure 4(a)) to represent relationships between patients. Arrows indicated who infected whom, creating a model of infection dynamics.

Pattern recognition played a critical role. They identified individuals with an empty "infected by" cell and explored cases like 1-8, which Manuel flagged as unusual. They also maintained a list of the 52 patients, marking off those who did not infect anyone, and color-coding groups based on who infected them (Figure 3).

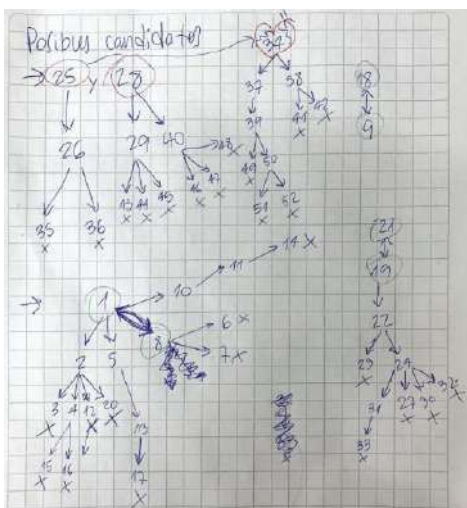
Identificador	Infectado por
Id0001	Id0008
Id0002	Id0001
Id0003	Id0002
Id0004	Id0001
Id0005	Id0001
Id0006	Id0008
Id0007	Id0008
Id0008	Id0001
Id0009	Id0018
Id0010	Id0001
Id0011	Id0010
Id0012	Id0002
Id0013	Id0005
Id0014	Id0011
Id0015	Id0004
Id0016	Id0004
Id0017	Id0013
Id0018	Id0009
Id0019	Id0021
Id0020	Id0002
Id0021	Id0019
Id0022	Id0019
Id0023	Id0022
Id0024	Id0022
Id0025	Id0025
Id0026	Id0025
Id0027	Id0024
Id0028	Id0028
Id0029	Id0028
Id0030	Id0024
Id0031	Id0024
Id0032	Id0021
Id0033	Id0031
Id0034	Id0025
Id0035	Id0026
Id0036	Id0026
Id0037	Id0034
Id0038	Id0034
Id0039	Id0037
Id0040	Id0028
Id0041	Id0038
Id0042	Id0038
Id0043	Id0029
Id0044	Id0029
Id0045	Id0029
Id0046	Id0040
Id0047	Id0040
Id0048	Id0040
Id0049	Id0039
Id0050	Id0039
Id0051	Id0050
Id0052	Id0050

Figure 3 Identification of who infected whom.

Finally, they recognized the need to repeat actions for suspected patient zeros, demonstrating pattern recognition within their algorithm.

Algorithm Design

The participants focused on achieving results before formalizing their approach. They first identified patient zeros, justifying their choices while excluding others. Though no explicit algorithm was stated early on, their actions show they developed one iteratively.



(a) Directed graph built by the participants.



(b) List created by Manuel.

Figure 4 Representations built to solve the problem.

Manuel excluded individuals who had not infected anyone using a list he systematically updated (as in Figure 4(b)). Meanwhile, Marcela used color-coding to visualize infection chains (as in Figure 3). Alejandra coordinated the information

provided by her teammates, asking systematically for each of the possible patient zeros identified initially and who infected them. At the same time, she built the directed graph (as in Figure 4(a)).

For everyone, Alejandra added nodes to the graph for each person they infected, marking them with their corresponding number and drawing arrows from the infector to the infected. If someone had not infected anyone, she marked their node with an X. She repeated this process for all nodes, moving on to the next potential patient zero until all candidates were reviewed. Finally, they verified the diagram against Manuel's initial list, marking off individuals as they were confirmed or excluded as patient zeros.

Once they completed the task, as the instructions explicitly required them to document their process, they wrote down the steps they followed. They worked sequentially to describe their actions, ensuring their written procedure matched the one they had completed moments earlier.

Debugging

Regarding debugging, the group corrected errors in the directed graph as they worked. For instance, they noticed during construction that they had omitted some individuals infected by 24 and initially misclassified 34 as patient zero. While the omission of 24's infections was corrected during graph construction, the error involving 34 was resolved upon revisiting the instructions.

After completing the task, they were asked about case 34. They explained that they had skipped verifying who infected 34 and instead jumped directly to identify whom 34 had infected. Alejandra admitted: "We didn't double-check everything for that one".

They also acknowledged making some reading errors while building the diagram, which they corrected along the way. Moreover, when writing their algorithm, they ensured the steps accurately reflected their actions (i.e. they realized they had performed an extra action earlier and adjusted one step to include it).

Iteration

Once the algorithm was written, no changes were made to it. Similarly, no significant adjustments were made to the initial procedure; instead, it was sequenced for clarity and to ensure it could be followed, despite some actions occurring in parallel. For example, Manuel's creation of a list to discard unchecked cases (originally intended for another purpose) and Marcela's use of color-coding were simultaneous. Thus, iteration does not appear to have been evident in this task.

Generalization

Participants were asked how they would handle a scenario with 100,000 individuals instead of 52. They responded that the same algorithm could be used, though without technological tools, they would need to "repeat, repeat, repeat many steps", increasing the likelihood of errors, as occurred with 52 cases.

To further assess whether their algorithm was generalizable, they were asked to consider using it in natural language, without programming knowledge or time constraints. In this context, they affirmed that their algorithm would still be applicable.

DISCUSSION AND CONCLUSION

The findings suggest that CT components (Shute et al., 2017) such as decomposition, abstraction, algorithm design, and debugging were evident in participants' actions. Generalization did not emerge naturally and required explicit prompting to extend the task to a larger numerical context. Iteration was not prominently observed, possibly due to time constraints or task demands. It could be encouraged through different task management, allowing students to revisit and test their algorithm with other data or optimize it.

Despite the participants' lack of formal training in CT or Discrete Mathematics, they addressed the problem using prior knowledge, demonstrating the task's potential to explore the relationship between MM and CT.

Unlike Lehmann's (2024) proposal, this study did not require students to explicitly follow MM phases. Nonetheless, evidence shows they transitioned between phases and engaged with cognitive processes described by Stillman et al. (2007), highlighting an area for further research.

Finally, it can be concluded that the preservice teachers successfully solved the MM task, which was designed to identify CT components. They developed a generalizable algorithm using representations such as tables and directed graphs. These findings align with previous studies and highlight the potential of MM tasks as effective tools for fostering CT development and underscores the importance of designing activities and research that further explore the relationship between Computational Thinking and Mathematical Modelling.

Acknowledgements: Beca Doctorado Nacional ANID 21241086, 21230712, 21220556.

References

- Abelson, H., y Kong, S. C. (Eds.). (2024). Computational Thinking Curricula in K–12: International Implementations. MIT Press.
- Ang, K.C. (2021). Computational Thinking and Mathematical Modelling. In: Leung, F.K.S., Stillman, G.A., Kaiser, G., Wong, K.L. (eds) Mathematical Modelling Education in East and West. International Perspectives on the Teaching and Learning of Mathematical Modelling. Springer, Cham. https://doi-org.pucv.idm.oclc.org/10.1007/978-3-030-66996-6_2
- Bardin, L. (2002). Análisis de contenido. Ediciones Akal.
- Bocconi, S., Chiocciariello, A., Kampylis, P., Dagienė, V., Wastiau, P., Engelhardt, K., Earp, J., Horvath, M., Jasutė, E., Malagoli, C., Masiulionytė-Dagienė, V., & Stupurienė, G.

- (n.d.). State of play and practices from computing education REVIEWING COMPUTATIONAL THINKING IN COMPULSORY EDUCATION.
- Bravo-Preciado, W., Castiblanco-Carrasco, R. A., & Pascuas-Rengifo, Y. S. (2024). Perspectiva crítica para la enseñanza del pensamiento computacional. *Revista Politécnica*, 20(39), 196–207. <https://doi.org/10.33571/rpolitec.v20n39a14>
- Kaiser, G., Blum, W., Borromeo Ferri, R., & Greefrath, G. (2023). Mathematisches Modellieren. In *Handbuch der Mathematikdidaktik* (pp. 399–428). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-66604-3_13
- Greefrath, G., Siller, H.-S., Vorhölter, K., & Kaiser, G. (2022). Mathematical modelling and discrete mathematics: opportunities for modern mathematics teaching. *ZDM – Mathematics Education*, 54(4), 865–879. <https://doi.org/10.1007/s11858-022-01339-5>
- Niss, M. y Blum, W. (2020). *The Learning and Teaching of Mathematical Modelling*. Routledge.
- Lehmann, T. H. (2024). Mathematical modelling as a vehicle for eliciting algorithmic thinking. *Educational Studies in Mathematics*, 115(2), 151–176. <https://doi.org/10.1007/s10649-023-10275-4>
- Leung, F. K. S., Stillman, G. A., Kaiser, G., & Wong, K. L. (Eds.). (2021). *Mathematical modelling education in East and West*. Springer.
- Peracaula-Bosch, M., & González-Martínez, J. (2024). Towards a Hermeneutics of Computational Thinking in Wing's Approximations. *Journal of Educational Computing Research*, 61(8), 139-158.
- Schukajlow, S., Kaiser, G., & Stillman, G. (2023). Modeling from a cognitive perspective: theoretical considerations and empirical contributions. *Mathematical Thinking and Learning*, 25(3), 259–269. <https://doi.org/10.1080/10986065.2021.2012631>
- Shute, V. J., Sun, C., & Asbell-Clarke, J. (2017). Demystifying computational thinking. *Educational Research Review*, 22, 142–158. <https://doi.org/10.1016/j.edurev.2017.09.003>
- Stake, R. (2015). *Investigación con estudio de casos*. Morata.
- Stillman, G., Galbraith, P., Brown, J., & Edwards, I. (2007). A framework for success in implementing mathematical modelling in the secondary classroom [Paper presentation]. 30th Annual Conference of the Mathematics Education Research Group of Australasia, Hobart, TAS, Australia.
- Weintrop, D., Beheshti, E., Horn, M., Orton, K., Jona, K., Trouille, L., & Wilensky, U. (2016). Defining Computational Thinking for Mathematics and Science Classrooms. *Journal of Science Education and Technology*, 25(1), 127–147. <https://doi.org/10.1007/s10956-015-9581-5>
- Wing, J. M. (2006). Computational thinking. *Communications of the ACM*, 49(3), 33–35. <https://doi.org/10.1145/1118178.1118215>
- Wing, J. M. (2014). Computational Thinking benefits society. *Social Issues in Computing* (blog). <http://socialissues.cs.toronto.edu/index.html%3Fp=279.html>

PROSPECTIVE SECONDARY MATHEMATICS TEACHERS' USE OF INQUIRY-BASED TEACHING PRINCIPLES TO PLAN A LESSON

Paloma Santonja, Mar Moreno, Salvador Llinares

University of Alicante

The aim of this research is to characterize prospective secondary mathematics teachers' use of Inquiry-based teaching principles when planning a lesson as a way of representing into practice the 5E instructional model. Lesson plans delivered by 25 prospective mathematics teachers (PMTs) were analysed considering both the tasks and teachers' anticipated movements. Results show two ways of using inquiry-based teaching principles to represent the 5E instructional model: (1) As isolated principles, and (2) as an articulated set of principles making a conceptual whole. Results show that only when the set of inquiry-based teaching principles are considered as a conceptual whole, prospective teacher are able to plan lessons reflecting the 5E instructional model conducting to exploratory learning.

INTRODUCTION

The Spanish curricular reform encourages teachers to design mathematics lessons in order to develop secondary students' mathematical thinking, for example, through reflections and establishment of conjectures. This way of promoting mathematical thinking requires adopting student-centred approaches, being inquiry-based learning (IBL) one of the recommended ones (Maaß & Artigue, 2013), which encourages students to think and act as mathematicians. The 5E instructional model provides a means of implementing this approach (Engage- Explore- Explain- Elaborate and Evaluate) (Bybee et al., 2006).

Literature review has shown the effectiveness of the 5E instructional model for enhancing secondary students' mathematical thinking when inquiry-based teaching principles are put into practice (Omotayo et al., 2017, Ünlüer & Kurtuluş, 2021). However, research has also addressed difficulties faced by teachers when applying this instructional model in mathematics education (Turan & Matteson, 2021). These authors have shown that mathematics teachers employing inquiry-based teaching methods often struggle to fully engage students. This difficulty arises from challenges in developing activities aligned with student interests or fostering curiosity about mathematical concepts, thereby hindering collaborative learning activities such as observation, investigation, analysis, and conclusion-drawing. This kind of difficulties reflects the need for developing formative programs in order to help teachers and prospective teachers learn to implement inquiry-based teaching principles (Calleja et al., 2024; Moreno et al., 2024). For example, Moreno et al. (2024) investigated how prospective secondary mathematics teachers used inquiry-based teaching principles

when they modified tasks. They found that PMTs could use these principles consistently only when specific mathematical practices were considered as learning goals. Additionally, Calleja et al. (2024) investigated how secondary mathematics teachers structured their inquiry lessons and examined how and why they altered the initial structure. Their results showed teachers departed to design three-phase lessons formed by presentation, work in small groups and conclusion; and after experiencing with the methodology, they were able to pose small subtasks and guide better the exploration.

However, literature review shows there is still need to investigate about how prospective mathematics teachers learn to use the inquiry-based learning principles when planning a lesson and how they adapted these principles to the contextual constraints. In this paper, we aim to contribute to fill this gap in the literature.

THEORETICAL FRAMEWORK

Traditional teaching is often characterized as a teacher-centred approach where the instructor (the teacher) explains concepts, provides examples to students, and assigns application tasks for later correction (e.g., Omotayo & Adeleke, 2017). In contrast, inquiry-based teaching is student-centred model, with the teacher's role encompassing (Maaß & Artigue, 2013, p. 782):

Orienting students towards questions and problems of interest for them that contain interesting learning potential; making constructive use of students' prior knowledge; supporting and guiding when necessary their autonomous work; managing small group and whole class discussions; encouraging the discussion of alternative viewpoints; and helping students to make connections between their ideas and relate these to important mathematical and scientific concepts and methods.

For putting the inquiry-based teaching principles into the practice the 5E instructional model sets five phases (Engage- Explore- Explain- Elaborate- Evaluate) through which students may work collaboratively to observe, investigate, analyse, and draw conclusions (Bybee et al., 2006). In the engagement phase, teacher can assess learners' prior knowledge and pose short activities that promote curiosity. In the exploration phase, teacher provides students with activities for generating new ideas upon their prior knowledge and explore questions. The activities demand a problem or question for stating mathematical properties or concepts is explored, looking for properties in sets of particular cases, conjecturing and checking conjectures in new sets of particular cases, searching for generalization and statements, proving or searching for counterexamples. The teacher's role in this phase consists of guiding the exploration by giving clues or asking questions for orienting students in their investigation. In the explanation phase, teacher must drive students' attention to particular aspects of the exploration. Firstly, letting students share their conclusions and secondly formalising mathematical concepts/properties. In the elaboration phase, students must develop a deeper comprehension or skills through additional activities, extending the previous tasks or proposing new ones to deepen in the comprehension of the mathematical

concepts. In the evaluation phase, the teacher can propose formal evaluative tasks for assessing the knowledge.

The 5E instructional model incorporates inquiry-based teaching principles (Maaß & Artigue, 2013) as follows: 1) The lesson begins with a teacher-introduced problem, question, or experiment (Engagement); 2) Students then undertake tasks demanding cognitive exploration of the problem or question, formulating mathematical properties or concepts under teacher guidance (Exploration); 3) The teacher subsequently teaches content based on student responses (Explanation); 4) The teacher facilitates generalization, extending or proposing tasks to deepen understanding of mathematical concepts (Elaboration); and 5) Finally, the teacher may use evaluative tasks to assess student learning (Evaluation).

OBJECTIVE AND RESEARCH QUESTION

This study investigates how prospective secondary mathematics teachers (PMTs) apply inquiry-based teaching principles during lesson planning. For that, we try to answer this research question:

How do prospective secondary mathematics teachers use inquiry-based teaching principles in their lesson planning decisions?

METHOD

Twenty-five PMTs enrolled in a one-year "Teaching Mathematics" course at the University of Alicante, Spain, participated in this study. The course, part of their teacher training program, involved 50 hours of face-to-face workshops over 15 weeks. These workshops facilitated reflection on curricular constraints and inquiry-based teaching principles, exploration of adaptable mathematical tasks, and consideration of innovative teaching practices. The course emphasized the development of teaching competencies, including lesson planning and the management of mathematical classroom discussions. For that, PMTs also engaged in other activities, including reading about: (i) curricular standards, (ii) inquiry-based teaching principles (Artigue and Blomhøj, 2013), and (iii) practices for orchestrating productive discussions (Stein et al., 2008). As part of the final assessment, PMTs were asked to design a lesson plan for one of these mathematical concepts: derivate, linear systems, semblance and progressions; with the goal of developing Inquiry-based teaching principles and integrating GeoGebra. PMTs were provided with a template (Figure 1) to solve this professional task.

PLAN: Plan a lesson for one week (4 sessions, that integrate TIC) for developing the chosen content from an Inquiry-based teaching perspective. Indicate: main goals, tasks sequence, specific aspects of inquiry-based teaching, and assessment criteria. Justify your lesson plan supports inquiry-based teaching and role play by TIC's.

Figure 1: Template for designing the lesson plan

Data for the study were the 25 lesson plans delivered by the participants. Qualitative analysis was carried on in two phases. In phase one, lesson plans were analysed considering (i) task placement and sequencing within the lesson, and the intended learning goals, and (ii) how the management of the mathematical discussion in the lesson was considered (anticipated specific movements of teachers). Following an inductive analytical process of generation of categories, we generated five categories for the tasks and the goal aimed: application, exploration, verification, demonstration and assessment. *Application tasks* were those whose finality was to apply concepts/properties previously explained by the teacher or explored by students. *Exploration tasks* were those whose finality was to investigate any mathematical property or concept, either by studying particular cases, searching for a kind of generalization or stating the falsity of a property or searching for counterexamples. *Verification tasks* were those in which students had to check mathematical properties/conditions following guidance in the statement. *Demonstration tasks* were those whose finality was to explicitly demonstrate a mathematical property. Finally, *assessment tasks* were those specifically to evaluate student progress.

Secondly, from an inductive analytical process, the aspects of the practice of teachers to manage the teaching and the mathematical discussions, we generated six categories: direct explanations, explanations with discussions, exemplifications, summaries, reviews, guiding an exploration using questions or clues. *Direct explanations* refer to instances where concepts are explained using a traditional teaching approach. *Explanations with discussions* were fragments referred to moments, after an exploration, in which PMTs elicit student conclusions to arrive at mathematical concepts or properties. Exemplifications were specific declarations of posing examples for clarification. Summaries were those fragments in which PMTs declared they would review the concepts explained in a lesson. Reviews were those fragments in which PMTs declared they would review concepts of the lesson until that moment. Finally, guiding an exploration were those fragments in which PMTs declared they would ask questions or give clues for monitoring the exploration, give clues for helping students deducing mathematical properties, facilitate the search of particular cases.

From this previous analysis, we framed the lesson plans considering to which extent they reflect the 5E instructional model as a way to put into practice the set the inquiry-based teaching principles. Initially, we organise the lesson plans into two non-exclusive groups: Lesson Plans reflecting the phases of the 5E model and lesson plans reflecting other aspects of instruction more centre into the teacher (traditional lessons). Table 1 displays the features of these two groups consider the teacher's anticipated movements and tasks in the different phases in the 5E instructional model. This framing let us identify ways of using inquiry-based teaching principles when planning a lesson.

		Tasks	Teachers' movements
5E instructional model	Engage	Exploration tasks	Reviews
	Explore	Exploration tasks	Guiding exploration
	Explain	Verification tasks	Explanations with students
	Elaborate	Demonstration tasks Exploration tasks as a continuum of previous ones	
Traditional teaching Phases	Evaluation	Assessment tasks	
	Direct Explanations		Direct explanation, Exemplifications, Summaries
	Application	Application tasks	
	Evaluation	Assessment tasks	

Table 1: Tasks and anticipated teacher's movements included in the lesson plan

RESULTS

Results show two ways in which PMTs used inquiry-based teaching principles when they meant to plan a lesson: (1) Using principles as isolated items (2) Articulating inquiry-based teaching principles as a conceptual whole. In this communication we describe these two ways of using the inquiry-based teaching principles.

Using inquiry-based teaching principles as isolated items

Some lesson plans included introductory problems, questions, or investigative tasks; however, these lacked a structured sequence where explanation followed exploration. Furthermore, the investigative tasks insufficiently supported key mathematical processes like examining specific cases, formulating conjectures, searching for counterexamples, or drawing conclusions. Teacher guidance was also less student-centred. For example, in PMT1's lesson plan on derivatives (for 16-17-year-old secondary students), a project on pandemic evolution was introduced in the first session, as indicated by:

Session 1: Presentation of the project and motivation of the students (10 minutes):
Introduce students the project Against COVID-19 with the help of a presentation developed at the beginning of the Learning situation.

While PMT1 utilized the principle of engaging students, the lesson plan primarily featured direct explanations and applications, neglecting other inquiry-based principles. Similarly, other PMTs incorporated tasks designed to investigate mathematical properties but omitted crucial mathematical practices such as generating

specific cases (observation), identifying patterns (analysis), or formulating conjectures (conclusion-drawing). PMT10's lesson plan on derivatives further illustrates this point. The plan involved the teacher explaining the derivative definition and then assigning a task requiring students to deduce derivative formulas: "Obtain the analytical expression of the derivative of the functions: (i) constant function, (ii) linear function, (iii) power function".

While this task allowed for formula derivation, it primarily focused on application of a pre-defined concept rather than encouraging exploration through case studies or independent investigation.

Articulating inquiry-based teaching principles as a conceptual whole

Some PMTs effectively integrated multiple inquiry-based teaching principles to structure their lessons. For instance, PMT2's lesson plan on deducing Thales' Theorem involved having students use GeoGebra to construct two secants and two parallel lines, then investigating relationships between the segments formed. This activity was designed to foster observation, investigation, analysis, and conclusion-drawing, with teacher guidance throughout the learning process:

Manipulate the configuration. Can you observe a relationship between AB , $A'B'$, BC and $B'C'$? What happens when dividing length of segment AB by length of segment $A'B'$? What about doing the same with the segments BC and $B'C'$? Does it happen always? Try and experiment.

Moreover, PMT2 indicated some teachers' movements to guide students while they were working in groups:

Teacher will move around the class, listening to students reasoning and making them questions, letting them continue with the resolution when needed. He will encourage trying with different particular sets (parallel and non-parallel lines) and explain their conclusions.

PMT2 planned students to expose their conclusions and teacher to make a summary:

Teacher will require students to explain their conclusions when using different procedures. The whole class will conclude that all the conclusions are equivalent and will write, guided by the teacher, the formulation of Thales' Theorem.

This fragment reflected the use of the inquiry-based teaching principle 3) *the teacher subsequently teaches content based on student responses*, related to the "Explanation phase" of the 5E-model. After that, the PMT2 planned students to demonstrate the theorem. The task encouraged students to look for a demonstration on their own and PTM2 anticipated some clues which students could use:

Let's demonstrate Thales' Theorem. Can you think of a way? If you don't know how to begin or continue, you can check the clues (we will use the notation of the previous task). What do we want to prove?

This task reflected the use of the principle 4) *the teacher facilitates generalization, extending or proposing tasks to deepen understanding of mathematical concepts*, as

students could get a deeper understanding of the Thales' Theorem by demonstrating it for a general configuration.

The application of inquiry-based teaching principles varied across lesson plans. While some plans incorporated the 5E instructional model's phases in only one or two sessions, others integrated the model throughout the entire lesson. These variations highlight the PMTs' ongoing development in applying inquiry-based principles to lesson planning.

CONCLUSIONS

The aim of this study was to characterize different ways in which prospective mathematics teachers use inquiry-based teaching principles when they plan a lesson. Findings show two ways of using inquiry-based teaching principles: as isolated principles, and as an articulated set of principles. Using this set of principles of an articulated way meaning that PMT are considering them in the specific order, like a conceptual whole, and linked to the possibility of that students observe, investigate, analyse and draw conclusions as specific mathematical practices. These ways of using inquiry-based teaching principles sheds light of how prospective teachers learn to plan lessons. However, a group of PMTs reflects a limit and no-articulate use of inquiry-based teaching principles bringing to the PMTs to design lesson plans which do not support the generation of opportunities to develop specific mathematical practices (Moreno et al., 2024), although they considered some specific characteristics such as engaging students. This finding extends Turan and Matesson's (2021) research on the absence of the Engagement phase in lesson plans. We concur that considering students' prior knowledge presents a greater challenge for prospective teachers than simply engaging students through real-world applications of mathematical concepts.

Our findings indicate that when PMTs make a use articulated of the inquiry-based teaching principles, their lesson plans exhibit cycles encompassing at least the Explore-Explain-Elaborate phases, differing from the structure found by Calleja et al. (2024) as departure point: Posing demanding tasks - Working on small groups - Mathematical discussions. We posit that the provided inquiry-based principles offer a valuable foundation for lesson planning, facilitating a shift from teacher-centred to student-centred approaches supportive of inquiry-based teaching. However, further research is needed to explore how PMTs apply these principles during actual lesson delivery and how teacher educators can best support their development in this area.

Acknowledgment

This research has been supported by the Agencia Estatal de Investigación, Ministerio de Ciencia e Innovación (Spain) under grants number PID2023-149624NB-I00; number PID020-116514GB-I00; and under the pre-doctoral contract grant PRE2021-097154.

References

- Artigue, M., & Blomhøj, M. (2013). Conceptualizing inquiry-based education in mathematics. *ZDM Mathematics Education*, 45(6), 797-810. <https://doi.org/10.1007/s11858-013-0506-6>
- Bybee, R. W., Taylor, A. J., Gardner, A., Van Scotter, P., Powell, J. C., Westbrook, A., & Landes, N. (2006). *The BSCS 5E instructional model: Origins, effectiveness, and applications, Full report*. Biological Science Curriculum Study. https://media.bsccs.org/bsccsmw/5es/bscs_5e_full_report.pdf
- Calleja, J., Foster, C., & Hodgen, J. (2024). Teachers' structuring of mathematical inquiry lessons: shifting from "task-first" to "scaffolded inquiry". *Research in Mathematics Education*, 26(3), 460-493. <https://doi.org/10.1080/14794802.2023.2176915>
- Maaß, K., & Artigue, M. Implementation of inquiry-based learning in day-to-day teaching: a synthesis. *ZDM Mathematics Education* 45, 779–795 (2013). <https://doi.org/10.1007/s11858-013-0528-0>
- Moreno, M., Llinares, S., & Santonja, P. (2024). Prospective secondary mathematics teachers' use of inquiry-based teaching principles as conceptual tools when modifying mathematical tasks. *Journal on Mathematics Education*, 15(4), 1129-1150. <http://doi.org/10.22342/jme.v15i4.pp1129-1150>
- Omotayo, S. A., & Adeleke, J. O. (2017). The 5E Instructional Model: A Constructivist Approach for Enhancing Students' Learning Outcomes in Mathematics. *Journal of the international society for teacher education*, 21(2), 15-26.
- Stein, M.K., Engle, R.A., Smith, M.S., & Hughes, E.K. (2008). Orchestrating Productive Mathematical Discussions: five Practices for Helping Teachers Move beyond Show and Tell. *Mathematical Thinking and Learning*, 10(4), 313-340. <https://doi.org/10.1080/10986060802229675>
- Turan, S. & Matteson, S. M. (2021). Middle school mathematics classrooms practice based on 5E instructional model. *International Journal of Education in Mathematics, Science and Technology*, 9(1), 22-39. <https://doi.org/10.46328/ijemst.1041>
- Ünlüer, İ., & Kurtuluş, A. (2021). The examination of conceptual and procedural understanding processes of eighth grade students in the subjects of identities and factoring. *Turkish Journal of Computer and Mathematics Education*, 12(1), 22-70. <https://doi.org/10.16949/turkbilmat.698535>

IT IS TIME: AN EYE-TRACKING STUDY ON CLOCK-READING

Maike Schindler¹, Mia-Milena Kölsch¹, Parviz Asghari², Anna Lisa Simon¹

¹University of Cologne, Cologne, Germany; ²TU Munich, Munich, Germany

Clock-reading is a significant aspect of students' lives and of the mathematics curriculum. School teachers are increasingly noticing that students are struggling with reading the clock, however, clock-reading has received little attention in research during the past decades. Our study aims to investigate how well school students are able to read the clock, and what strategies they use for clock-reading. Conducting an eye-tracking study with 74 ten-year-old students, we found that many students were struggling with clock-reading and that the success rates were substantially lower than in previous studies. Our study, using eye-tracking video analysis, also brought to light new insights into students' clock-reading strategies that expand the current state of research. At the same time, it shows the boundaries of eye-tracking data analysis.

INTRODUCTION

Measurement of time involves a complex conventional system that is not easy to understand for many children (Boulton-Lewis et al., 1997). However, using this conventional system by using calendars and reading the clock is an important goal for all students in primary school (Burny et al., 2012) and it is part of curricula in many countries (e.g., Department of Education, 2021). In particular, students need to learn to read and set both analogue and digital clocks (e.g., Ministry of School and Education NRW, 2021). Although digital clocks are being more and more used and present in students' everyday lives (Williams, 2012), analogue clocks still are present at important pivotal points (e.g., train stations, school classrooms, airports, public swimming pools, churches) and the analogue clock serves as model to understand times (Mutlu & Korkmaz, 2020). For example, to understand that 7:50 is ten to 8 or that 4:30 is half past 4, the analogue clock is a helpful visual model.

Although clock-reading is important for persons' everyday lives and is an important ability for students to learn, it has attracted little attention in mathematics education research (Burny et al., 2012; Mutlu & Korkmaz, 2020). Besides, the scarce findings about success rates and students' strategies reading the clock (e.g., Boulton-Lewis et al., 1997; Friedman & Laycock 1989) are more than 30 years old, and much has changed in students' living environments since then. For example, many students use smart phones that normally display digital rather than analogue clocks. Also, (smart) watches often use digital displays, while the use of wristwatches has generally decreased. The research gap on students' clock-reading skills calls for a current study investigating how well school students are able to read the clock, and what strategies they use for clock-reading. We ask the research questions, (1) *How well are students*

able to read the analogue clock, in particular, what are the success rates? and (2) What strategies do the students use to read the analogue clock?

To investigate how and how well the students can read the clock, we used eye tracking—the recording of students’ gazes (Strohmaier et al., 2020)—as a tool. While previous studies aiming to investigate students’ clock-reading strategies (e.g., Boulton-Lewis et al., 1997) drew on students’ self-reports, eye tracking offers potential for new insights into clock-reading strategies that are not constrained by students’ ability to express their thoughts (see Schindler & Lilienthal, 2018). In recent years, eye tracking has become a powerful tool for research in mathematics education due to technological and theoretical advancements (e.g., Strohmaier et al., 2020). Research involving visually presented tasks, such as number line and enumeration activities, has shown that eye tracking can be advantageous over students’ self-reports or thinking aloud for investigating student strategies, since it is less affected by students’ abilities and inclination to express themselves (Schindler & Lilienthal, 2018; Simon & Schindler, 2020). The current study—in addition to its empirical interest—inquiries into the possibilities that eye tracking holds for investigating students’ clock-reading strategies.

RELATED WORK

The dimension of time is one of the basic structures of human life. Already in the school entry phase, students should be able to name and set simple times (full hour, quarter hour) on analogue and digital clocks—and by the end of grade 4, students are expected to expand their skills to read and set more complex times on analogue and digital clocks (Ministry of School and Education NRW, 2021). Although digital clocks are nowadays more common than analogue ones (especially through smart phones etc.) and their influence is increasing, still being able to use analogue clocks is important for the students’ lives as mature individuals; and the analogue clock serves as a model for the students to develop a sense of time (Mutlu & Korkmaz, 2020). This is why learning the analogue clock comes first in many curricula (for England, see, e.g., Department of Education, 2021).

Clock-reading involves both absolute and relative time (Williams, 2012). For absolute time, the students typically read the hour first and the minutes afterwards. For example, the time in Fig. 1 in the top row, second from left would be read as “seven fifteen”. In relative time however, the minutes, which are shown by the long hand of the analogue clock, are read first (Williams, 2012). In the 7:15 case, the time would be “quarter past seven”. Internationally and even intranationally, there are differences in conventional systems for relative time (Williams, 2012). In the part of Germany where this study was conducted, the conventions for relative times are as shown in Fig. 1.

Research investigating how well students can read the clock found that 8-year-old children can read full hour and half-hour times with high accuracy (97% to 100%), quarter-hour times with 80% accuracy, five-minute times with 79%, and minute times with 38% to 78% accuracy (Friedman & Laycock, 1989; Siegler &

McGilly, 1989). However, these numbers stem from studies conducted more than 30 years ago and the question arises if these numbers are still valid nowadays that smart phones and other digital tools displaying digital clocks are dominating our lives.

Regarding students' strategies to read the clock, Boulton-Lewis et al. (1997), Friedman and Laycock (1989), and Siegler and McGilly (1989) conducted studies in which they asked students from grades 4 to 6 to self-report their strategies when reading the clock. In Boulton-Lewis et al.'s study, the students reported that they, for example, (1) incremented in 5's and 1's, (2) they identified the short and long hand as hours and minutes, and (3) they referred to landmarks (quarters), saying, for example "3 is 15 minutes or a quarter past" (p. 144). These three strategies were the most common ones among the participating fourth, fifth, and sixth graders. However, these results stem from the 1990s and it needs to be investigated if students still use strategies in this way and prevalence. Additionally, by that time methods such as eye tracking were not yet available, and eye tracking promises to allow new insights into clock-reading activities.

THIS STUDY

Participants and setup. The sample of the present study consisted of a total of 74 students (mean age: 10.8 years, 44.6% female) from a fifth grade at a German inclusive comprehensive school. The study took place in a quiet room in the students' school.

Clock tasks. We used eight different times (Fig. 1) on an analogue clock. First, the students were shown a video explaining the format of *absolute* time. The students were then asked to read the absolute time for the eight times and say them out loud. The eight times appeared in randomized order, yet in the same order for every student. After this, another video was shown explaining the format of *relative* time. The students were then asked to read and say out loud the same eight times in the relative form in the same order as before. All students' oral answers were recorded manually/written down

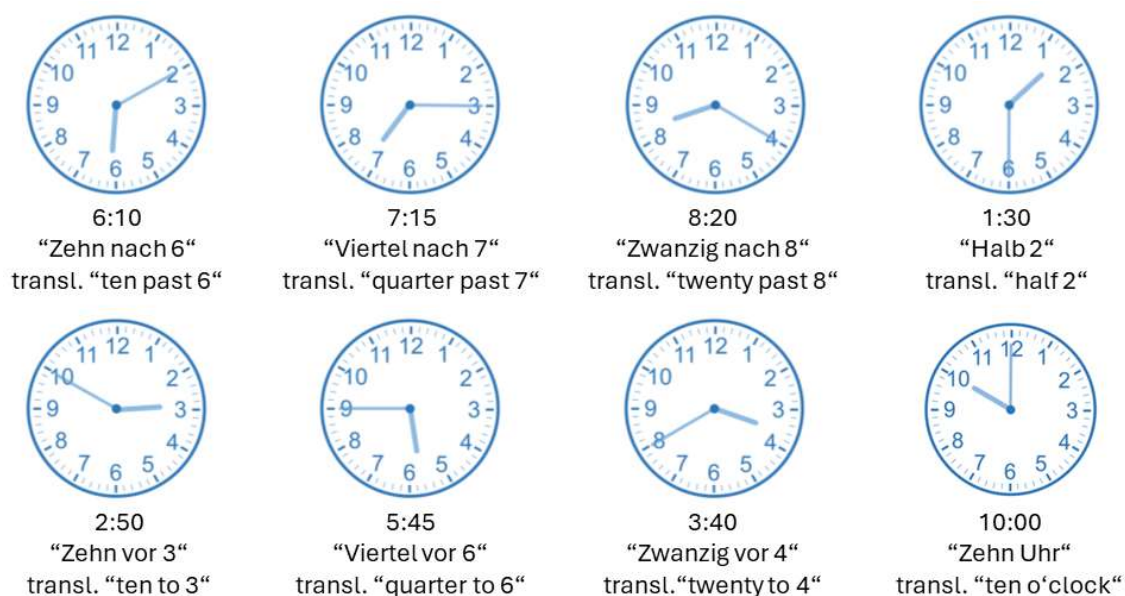


Fig. 1: Times used in this study with translations for relative time conventions

on an answer sheet, which resulted in 16 student answers recorded for each student.

Eye tracking. For the recording of eye movements, the Tobii Pro X3-120 eye tracker (120Hz, binocular, infrared) was used. This eye tracker was attached to a 24" full HD computer screen on which the tasks were displayed. The distance between the students and the screen was approximately 60cm. The system was calibrated for every student. The average accuracy of the eye-tracking data in the present study was 0.9°.

Data analysis. For the analysis of students' *answers*, we determined for each answer if it was correct, incorrect, or not in accordance with the conventions. The latter means that the students did not use quarters or half hours, but said, for instance, "30 past" instead of "half ...", or they said, "10 past half" instead of "20 to". We then calculated and graphically visualized the percentages of the correctly read times. To analyze students' *clock-reading strategies*, eye-tracking videos in which the gaze was displayed as semi-transparent dot were used. Qualitative content analysis according to Mayring (2000) was used to analyze the strategies—in an inductive manner, as the categories were derived directly from the material. First, all eye-tracking videos were watched, and the students' eye movements were paraphrased for each task separately. Subsequently, a codebook was created in which the reading of the hours and the reading of the minutes were defined. Subcategories with different observed strategies were defined for both the reading of the hour and the reading of the minutes. They were subsequently merged into four main categories, which are reported in this paper. The strategies were provided with examples of students' gazes in the form of a gazeplot, i.e., scanpath visualizations of students' eye movements (see Fig. 3 for examples). For the analysis of students' strategies, we decided to only include the data from correctly solved tasks, i.e., times where the students gave correct answers. This was because in cases when the students could not figure out the times, they often looked around on the clock face for a long time and at a multitude of things, partially repeatedly, so that a meaningful manual analysis of these videos could not be achieved. Other students who gave incorrect answers guessed and gave random answers, for which eye-tracking interpretation was not appropriate either (see reflection in Discussion section).

RESULTS

We found that the overall correctness was relatively low in general, with 63.85% correct solutions for absolute times and 35.98% correct solutions for relative times (Fig. 2). Non-conventional answers occurred in both absolute and relative times, but only in less than 3% for absolute times (here, the students, e.g., used relative instead of absolute times). Regarding the individual tasks/times, we found that while 10:00 was the time with highest correctness in both absolute and relative time, the correctness tendentially decreased with the number of minutes that the minute hand showed, with lowest correctness for 2:50 in absolute time and 3:40 (twenty to 4) in relative time.

When analyzing the eye-tracking videos to identify student strategies, we found four main categories (Fig. 3): (1) Direct orientation, (2) Counting in increments of 5, (3)

Use of orientation points (landmarks), and (4) Orientation towards mirror image. Figure 3 details these strategies and illustrates them with gaze plot examples, which visualize the gazes of single students when working on the task: The green dots indicate gaze fixations, and their order is indicated by the numbers given in the dots.

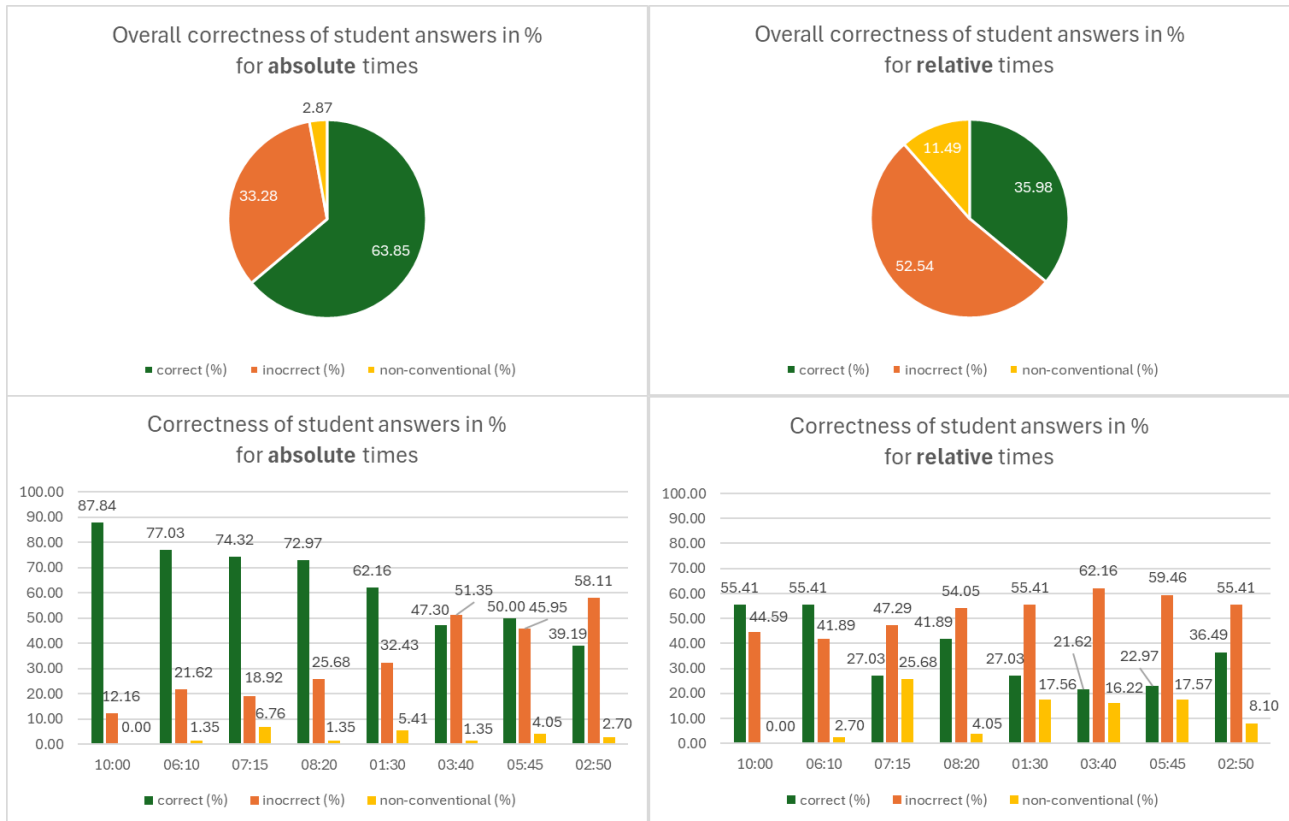


Fig 2: Correctness of student answers for absolute and relative times

The four main categories of student strategies apply to both students' reading of hours and minutes on the clock. However, for reading the hours the students used direct orientation in 99% of the cases both for absolute and relative time. For telling the minutes (Fig. 4), the students used direction in approx. two thirds of the cases in both absolute and relative times. For absolute times, they used orientation points roughly as

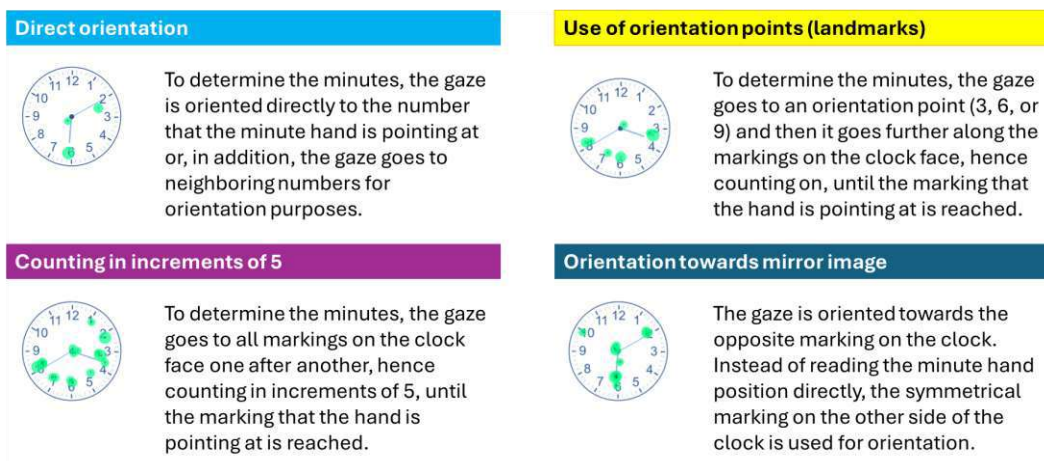


Fig 3: Main categories of student strategies to read minutes on the clock

often as incrementing in fives, while for relative times, the use of orientation points occurred slightly more often than counting in fives. Orientation towards mirror images was used rarely and only for relative times.

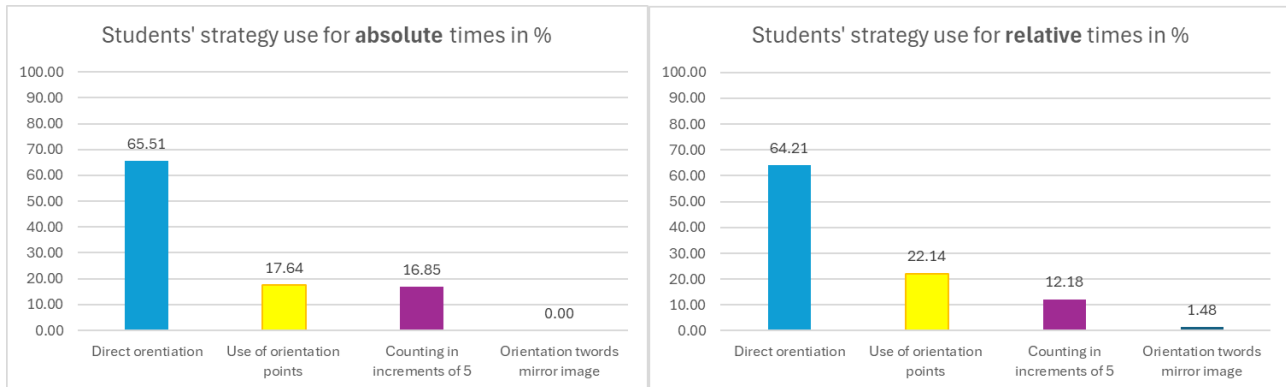


Fig 4: Strategy use for reading the minutes for absolute and relative times

DISCUSSION

This study aimed to investigate how well school students are able to read the clock, and what strategies they use for clock-reading. Conducting a study with 74 fifth-grade students (avg. 10.8 years old) and using different absolute and relative time reading tasks, we were able to determine students' success rates in reading the clock and we found out students' strategies and how often these were used by the students.

Regarding students' success in reading the clock, we found that the participating fifth graders read 64% of times correctly for absolute time and 36% correctly for relative time tasks. This was low compared to previous studies, which investigated how well 8-year-old students in the 1980s and 1990s read the time (Friedman & Laycock, 1989; Siegler & McGilly, 1989). While those studies found success rates of 97% to 100% for full hour and half-hour times, of 80% for quarter-hour times, and 79% for five-minute times, the success rates in our study were substantially lower, especially for relative time (e.g., success rate of 55% for 6:10), but also for absolute time (e.g., success rate of 50% for 5:45). This finding confirms the impression that many schoolteachers and parents have: That many students are nowadays struggling with reading the time on the analogue clock; struggles that relate to a change in students' living environments (e.g., with an increase of digital watches on the mobile phone). However, the analogue clock is still an important model to understand the daytime, also for understanding and interpreting time on digital clocks (see above), which is why learning to make sense of the analogue clock remains an important goal for students' lives.

The numbers in this study (Fig. 2) furthermore indicate that especially reading the minute hand is related to the difficulty level of the tasks: As a rough trend, the more minutes indicated by the minute hand, the lower the success rate. This is plausible since one can tell the hour relatively straightforwardly from the clock face. For example, for 6:10 the hour hand points towards the 6 on the clock face (while the minute hand points to 2, which needs to be interpreted as 10 min). In our analysis of students' strategies,

this was confirmed in that the hour hand was perceived straightforwardly (direct orientation) in 99% of the cases. Nevertheless, for times like 2:50, where the hour hand is closer to 3 than to 2, mistakes can occur when students read 3:50 instead of 2:50.

When investigating students' time reading strategies, we methodologically found that for tasks where the students could not figure out the correct time, the eye-tracking data were often either too complex or not meaningful for a manual eye-tracking video analysis. In our study, we did not use thinking aloud, since we wanted to avoid an influence of students' verbalization on their strategies, however, in these cases, additional thinking aloud (possibly retrospective) could be beneficial in future studies. Also, automated eye-tracking data analysis using AI could be beneficial for the future to make sense of the eye-tracking data of incorrectly solved tasks, for instance, to find clusters within these data that might be interpretable for human researchers—in a similar vein as performed by, for example, Schindler et al. (2022).

Through eye tracking, we were able to find four main categories (and several subcategories) of student strategies for reading the clock: (1) Direct orientation, (2) Using orientation points (quarter past, half, quarter to), (3) Counting increments of five, and (4) Orientation towards mirror image. These categories had not been found in this way before, also since previous studies used students' self-reports and did not have access to eye tracking. However, the categories we found partially relate to strategies that students reported on in previous studies, which makes our findings compatible with previous results. For example, in Boulton-Lewis et al.'s (1997) study students reported that they also had incremented in 5s or 1s. Also, Boulton-Lewis et al. (1997) found the strategy "landmarks" where, for example, "the child said 3 is 15 minutes or a quarter past" (p. 144), which might relate to the use of orientation points in our study. Regarding the prevalences, we found that for reading the minute hand, the students used direct orientation most often, in approx. two thirds of the cases, followed by the use of orientation points with approx. a fifth of the cases. Counting in increments of fives occurred only in 17% (absolute time) and 12% (relative time) of the cases, which is an interesting finding given that in Boulton-Lewis' et al.'s (1997) study, 53% of the fifth graders reported that they incremented in 5s/1s, which was the strategy used most often. This difference might be caused by the fact that our analysis focused on strategies for correctly solved tasks, and that succeeding students may have a greater tendency to recall images of the clock and "see it", while incrementing in 5s might be more prevalent among struggling students. Also, the difference might be caused by the different data: In self-reports, students might have the tendency to report counting even when they directly "see" it when they feel it is not adequate to say they just saw it.

To sum up, our study found that the fifth graders made more mistakes in clock-reading tasks than it was reported in earlier studies, which may be related to the students' changes in living environments during the past decades. This shows that it is necessary to investigate students' clock-reading skills, which are significant despite the increased digitalization, and that clock-reading needs educational attention, for instance, in form

of intervention studies. Our study brought forward new categories of clock-reading strategies that had not been found before, which was possible through the use of eye tracking. Eye-tracking video analysis was only viable for a part of students' clock-reading activities, yet it provided new insights, which expand the current state of research.

References

- Boulton-Lewis, G., Wilss, L., & Mutch, S. (1997). Analysis of primary school children's abilities and strategies for reading and recording time from analogue and digital clocks. *Mathematics Education Research Journal*, 9, 136–151.
- Burny, E., Valcke, M., & Desoete, A. (2012). Clock reading: An underestimated topic in children with mathematics difficulties. *Journal of Learning Disabilities*, 45(4), 351–360.
- Department of Education (2021). *National curriculum in England: mathematics programmes of study*. <https://www.gov.uk/government/publications/national-curriculum-in-england-mathematics-programmes-of-study/national-curriculum-in-england-mathematics-programmes-of-study>
- Friedman, W. J. & Laycock, F. (1989). Children's analog and digital clock knowledge. *Child Development*, 60, 357–371.
- Mayring, P. (2000). Qualitative content analysis. *Forum: Qualitative Social Research*, 1(2), Art. 20.
- Ministerium für Schule und Bildung des Landes NRW (2021). Lehrpläne für die Primarstufe in Nordrhein-Westfalen. <https://www.schulentwicklung.nrw.de/lehrplaene/>
- Mutlu, Y. & Korkmaz, E. (2020). Investigating clock reading skills of third graders with and without dyscalculia risk. *International Online Journal of Primary Education*, 9, 97–110.
- Schindler, M., Doderer, J. H., Simon, A. L., Schaffernicht, E., Lilienthal, A. J., & Schäfer, K. (2022). Small number enumeration processes of deaf or hard-of-hearing students: A study using eye tracking and Artificial Intelligence. *Frontiers in Psychology*, 13:909775.
- Schindler, M. & Lilienthal, A.J. (2018). Eye-tracking for studying mathematical difficulties—also in inclusive settings. In E. Bergqvist, M. Österholm, C. Granberg, & L. Sumpter (Eds.), *Proc. Of the 42nd Conference of the IGPME* (Vol. 4, pp. 115–122). PME.
- Siegler, R. S. & McGilly, K. (1989). Strategy choices in children's time-telling. In I. Levin & Zakay (Eds.), *Time and human cognition: A life-span perspective* (pp. 185–218). Elsevier.
- Simon, A.L., & Schindler, M. (2020). A comparative analysis of eye tracking and thinking aloud in number line estimation tasks: A study on students with and without mathematical difficulties. In M. Inprasitha, N. Changsri, & N. Boonsena (Hrsg.), *Proc. of the 44th Conference of the IGPME* (Vol. 4, pp. 65–72). PME.
- Strohmaier, A. R., MacKay, K. J., Obersteiner, A., & Reiss, K. M. (2020). Eye-tracking methodology in mathematics education research: A systematic literature review. *Educational Studies in Mathematics*, 104, 147–200.
- Williams, R. F. (2012). Image schemas in clock-reading: Latent errors and emerging expertise. *Journal of the Learning Sciences*, 21(2), 216–246.

REAL-WORLD CONNECTIONS IN MATHEMATICS EDUCATION FOR STUDENTS WITH INTELLECTUAL DISABILITIES: A SYSTEMATIC LITERATURE REVIEW

Laura Schmidt¹, Janina Krawitz¹, Susanne Schnepel²

¹ University of Cologne, Germany; ² University of Münster, Germany

In this systematic literature review, we examined the most recent research (2020-2024) on real-world connections in mathematics education for students with special educational needs, particularly those with intellectual disabilities. Our analysis of 44 empirical studies revealed that this research primarily used experimental single-case designs, with most research conducted in the US. The majority of the studies employed standard word problems, while realistic word problems were rarely found, and only one study utilized authentic and open problems. The findings call for more consistent terminology in describing special educational needs and for the use of more authentic tasks to better help learners with special educational needs develop practical mathematical skills.

INTRODUCTION

It is important to make real-world connections in mathematics instruction to help students with special educational needs transfer mathematical content to real life (Browder et al., 2017). Such connections can help prepare them for an independent life in which they can use the skills they have learned (e.g., in a work environment) (Spooner et al., 2017). Thus, it is crucial to obtain valuable information on what these students need for successful instruction in real-world problems, especially in the general classroom, as students with special educational needs have the right to education within the general education system and should not be excluded due to their disabilities (United Nations, 2007).

A fundamental area in mathematics education is the application of mathematics to the real extra-mathematical world (Niss & Blum, 2020). Many questions and problems in students' mathematical learning are influenced by interactions between mathematics itself and the world in which students operate. The integration of real-world connections (e.g., in the form of word problems) can motivate students and help them develop mathematical problem-solving skills and techniques while simultaneously introducing mathematical concepts (Verschaffel et al., 2000).

In our literature review, we focused specifically on students with intellectual disabilities. Despite the recognized importance of real-world connections in mathematics education for these learners, no prior literature reviews have addressed this area. In conducting this literature review, our aim was to analyze the most recent research to identify the different types of special educational needs and real-world connections that have been explored as well as to highlight existing research gaps.

THEORETICAL BACKGROUND

Students with intellectual disabilities and their learning characteristics

Students with intellectual disabilities (ID) exhibit significant limitations in intellectual functioning and adaptive behavior (Schalock et al., 2021). The ICD-11 categorizes ID under “neurodevelopmental disorders” (WHO, 2019). The term “intellectual disabilities” serves as an umbrella term for various developmental disorders, learning disabilities, and conditions that are associated with specific syndromes. ID often include students at risk of academic failure or social exclusion and may include those with undiagnosed learning difficulties, socio-economic disadvantages, or inadequate access to support services, making early intervention critical. The distinction between students with ID and those at risk of developing them has yet to be clearly defined. Impairments in this cognitive area originate during the developmental period (Schalock et al., 2021). Students often face challenges in specific intellectual functions, such as reasoning, working memory, processing speed, and verbal comprehension (WHO, 2019). Additionally, they may exhibit specific learning impairments that affect reading, written expression, or mathematics. Other related impairments might involve speech and language difficulties or co-occurring conditions such as Attention Deficit Hyperactivity Disorder (ADHD). In many cases, these challenges overlap, reflecting the highly heterogeneous nature of this group (WHO, 2019). These students may be taught in specialized schools or in inclusive settings in mainstream education.

Real-world connections in mathematics education

The most common approach to integrating mathematics with the real world in lessons is through the use of problems that are based in reality. To solve such problems, processes for translating the problems from real-world contexts into mathematics and the other way around are necessary (Niss & Blum, 2020). Problems with real-world connections can serve two distinct purposes: as content or as a vehicle (Julie, 2019). When used as content, they are aimed at developing students’ skills in applying mathematics to real-world situations. By contrast, when used as a vehicle, they serve as a tool for learning mathematical content. Furthermore, mathematical problems originate within the domain of mathematics and are framed in realistic contexts, or else they originate from authentic real-world situations and can be solved using mathematical approaches (Niss & Blum, 2020). Accordingly, two types of mathematical tasks that are in some way connected to the real world can be distinguished: standard word problems and realistic word problems. Standard word problems are characterized by verbal descriptions of problems that include numerical information and question(s) prompting students to solve them using mathematical operations (Verschaffel et al., 2000). Typically, the operations and tools required to solve standard word problems are straightforward (Verschaffel et al., 2020), and the real-world context is often greatly simplified and dressed-up (Niss & Blum, 2020). By contrast, realistic word problems are more authentic and require more complex translations between the real world and mathematics (Kaiser, 2017).

Real-world connections in special needs education

Historically, special needs education has emphasized that mathematics should be taught through practical, real-life activities. In their literature review, Bowman et al. (2019) found many studies that focused on how students with special educational needs solve standard word problems. A meta-analysis of single-case research in word-problem solving for students of all ages with learning disabilities by Shin et al. (2021) found that specific word-problem instruction has a positive effect on students' abilities to solve such tasks. In another meta-analysis, Myers et al. (2022) found that additional support in the form of interventions can improve the performances of students who experienced mathematics difficulties in primary grades when solving word problems, and teachers can use existing interventions and apply heuristic strategies and approaches. However, realistic word problems have not been focused on as frequently in special needs education. Scott-Wilson et al. (2017) showed that including realistic word problems in mathematics lessons can enhance important life skills of learners with special educational needs and support their learning of mathematical concepts when tasks are appropriately differentiated to match their ability levels.

RESEARCH QUESTIONS

The current systematic literature review focused on the following research questions:

RQ1: What are the general characteristics (geographical distribution, sample size, school context, and research methods) of recent research on real-world connections in mathematics education for learners with special educational needs?

RQ2: Which special needs are addressed in studies in this research area?

RQ3: What types of problems with connections to the real world are addressed in this research area?

METHOD

Literature search and selection of studies

We decided to limit our review to empirical research published in journals between 2020 to 2024, as we consider them to be an appropriate representation of most current research in our area of interest. On September 17, 2024, we conducted a search in the databases Web of Science Core Collection, Scopus, and ERIC using the search terms math*, model*, application*, word problem*, real*, special education*, special need*, intellectual disabilit*, learning disabilit*, developmental disabilit*, and cognitive disabilit*. The search was limited to the title, abstract, and keywords and was only related to peer-reviewed articles written in English.

From this search, we identified 707 articles (497 without duplicates) that might be relevant for this study. To be included in this review, articles had to fulfill the following criteria: the research topic is located in educational research (1a) or more precisely mathematical educational research (1b). The study addresses real-world connections (2a) (e.g., word problems, mathematical applications, or mathematical modelling), or

real-world connections are the main focus of the study (2b). Articles consider students with special educational needs (3) and are in the area of empirical research, excluding meta-analyses and reviews (4). The opposite conditions were applied as exclusion criteria (EC). Figure 1 presents the entire study identification procedure.

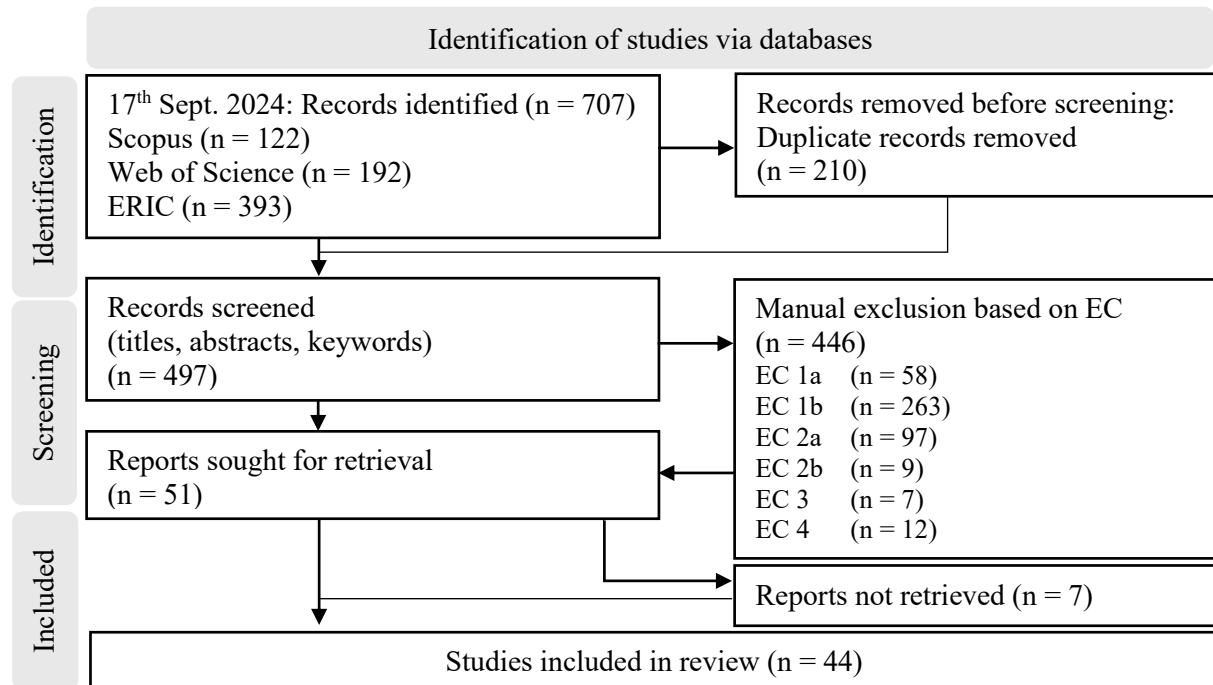


Fig.1: Study identification procedure

Data extraction and analysis

We conducted a qualitative content analysis of the full texts to systematically extract relevant data. Thus, we created a coding scheme to categorize the 44 articles in accordance with the research questions. By screening the full texts, we first coded (RQ1) the geographical distribution, the sample size (categorized into ranges: 1-3, 4-10, 11-20, 21-50, 51-100, 101-500, and 501-1000), the school setting, and the research design (single-case, experimental, and descriptive research designs).

Furthermore, we coded different types of special educational needs (RQ2). Regarding the special educational needs, we first categorized the terms used by the authors inductively. For this purpose, each special need mentioned was coded once per study, resulting in several special educational needs being assigned to each study. Afterwards, we divided these terms into three overarching categories. The first category, *development and cognition*, included intellectual and developmental disabilities. The second category, *learning impairments*, referred to students at risk of learning difficulties and those with learning difficulties and mathematical or specific learning disabilities such as dyscalculia. In addition, the last category, *other or co-occurring impairments*, covered all additional and other impairments mentioned by the authors.

To assess the real-world connection (RQ3) for each study, we coded whether real-world connections were used as *content* or as a *vehicle* for teaching mathematical concepts. Furthermore, the terms the authors used to describe the nature of the real-

world connection in the tasks they used were coded inductively. Afterwards, these terms were analyzed and compared in order to define the categories. We thereby differentiated *between standard word problems* and *realistic word problems*. To check whether the coding was intersubjectively applicable, ten of the articles were coded by a second person with satisfying agreement (Cohen's $\kappa \geq .84$).

RESULTS

To address RQ1, we analyzed the general characteristics of the 44 studies. Participants were predominantly from schools in the US (30 studies, 68%), followed by Spain (4 studies, 9%), China (2 studies, 5%), and one study each (2%) from eight other countries (Belgium, Indonesia, Jordan, Kuwait, Saudi Arabia, South Africa, Taiwan, Turkey). The majority of studies had small sample sizes: 21 studies (48%) had between one and three participants, 9 studies (20%) included four to ten participants, and 4 studies (9%) had between eleven and twenty participants, as well as 4 studies (9%) between fifty-one and one hundred participants. From the remaining studies, 2 (5%) considered twenty-one to fifty participants and 4 (9%) included one hundred one up to 692 participants. A total of 37 studies (85%) were conducted with students from regular schools, while 7 studies (16%) were conducted in special schools. Thirty-eight studies (86%) used experimental designs, with 27 (61%) utilizing single-case designs. These included methods such as multiple-baseline designs and multiple-probe designs across different aspects. The remaining 11 experimental studies (25%) included 9 quasi experimental (20%) and 2 true experimental designs (5%). Quasi experimental designs show pre-posttest designs for one group or non-equivalent groups. True experimental designs include randomized groups and pre-posttest designs. Besides the thirty-eight experimental studies, the remaining 6 studies (14%) used descriptive research approaches, such as interviews, observations, or analyses of differences in students' performances, strategies, or explanations.

To address RQ2, we analyzed the types of special educational needs that the students participating in the studies had. More than half of the studies (24 studies, 54%) considered students with special educational needs regarding *development and cognition*. These studies included students with "autism spectrum disorders," "intellectual disabilities," "developmental disabilities," and "Down syndrome." Furthermore, 26 studies (59%) focused on students with *learning impairments*. These included "mathematical difficulties" or mathematical disabilities, such as "dyscalculia" as well as "slow learners" and students with "general learning disabilities" or "specific learning disabilities," although the last one was often not explicitly defined. The remaining studies focused on students with *other or additional occurring impairments* (16 studies, 37%). These included language and speech impairments, "attention deficit hyperactivity disorder," and "other health impairments."

To address RQ3, we analyzed the real-world connections in the selected studies. The vast majority of studies (35 studies, 80%) addressed real-world connections as a

vehicle. Real-world connections were predominantly used to teach mathematical concepts to students. Few studies (5 studies, 11%) addressed real-world connections as *content* for teaching the application of mathematics in the real world. In the remaining 4 studies (9%), real-world connections were not used as *content* or as a *vehicle* because only certain aspects of the process of solving such tasks were described or observed. Furthermore, most studies focused on *standard word problems* (39 studies, 89%). These included dressed up mathematical word problems such as “Mark has 11 blue and red crayons. If 7 of the crayons are red, how many are blue?” (Powell et al., 2020). In only a few studies (5 studies, 11%), the authors referred to *realistic word problems*, describing them as “word problems grounded in life skills,” “real-world thematic word problems,” or “real-world scenarios.” Analyzing the problems used in these studies revealed that four out of the five articles still used standard word problems. Only one study (2%) used authentic and open tasks and examined the implementation of these tasks in math lessons for students in the category learning impairments, specifically “slow learners.” For instance, a figure showing a coal mining area with different construction vehicles, cars, and a winding road was presented, and learners were asked “if it is possible for a car at the bottom to climb up to the surface with a straight vertical path” (Listiawati et al., 2023).

DISCUSSION AND IMPLICATIONS FOR RESEARCH

The aim of this review was to provide an overview of the most recent research on real-world connections in mathematical education for students with special educational needs specifically focusing on intellectual disabilities. The results highlight several main issues. The majority of the selected studies were conducted in the US. Since the structures within different countries have grown historically, and there are no uniform approaches, it is reasonable to assume that research on special needs education is often published in the respective national language (Grünke & Cavendish, 2016). Placing a stronger emphasis on international publishing would significantly advance the field of special needs education research. Furthermore, our results reveal a strong tendency toward small sample sizes. In special needs education, it can be a challenge to reach larger samples or to be able to draw general conclusions due to the great heterogeneity of students with special needs. Nevertheless, conducting studies with larger groups of participants could enhance the understanding of how real-world connections can be effectively implemented in math lessons for students with special educational needs.

Analyzing different special educational needs addressed in the studies revealed that research on real-world connections for learners with special needs has addressed various types of special educational needs, including those related to development and cognition, learning impairments, and co-occurring impairments. The inconsistent terms used by the authors (e.g., for learning impairments) might be explained by the fact that there are many different international approaches in the field of special needs education (Grünke & Cavendish, 2016). Although there are internationally accepted classification schemes, such as the ICD-11, they do not originate from the educational field, and they describe the special educational needs of students on a psychological or

health level, separate from school-related factors. As the individuality of the students plays a major role and should not be neglected, the question that arises is what an international approach might look like and whether it could be implemented. However, to foster knowledge accumulation and exchange in the field of special needs education, studies must place a stronger emphasis on the standardization of terminology as well as on describing students' mathematical preconditions (e.g., arithmetic skills, language comprehension, problem-solving skills).

Analyzing the real-world connections in the studies revealed a strong focus on standard word problems. Even though it is very important for students with special educational needs to learn life skills, independence, and autonomy in their everyday school life, the tasks that were used were rarely realistic. Therefore, it is important to find out how authentic and open realistic word problems can contribute to the mathematical education of students with special educational needs, as in the studies by Listiawati et al. (2023) and Scott-Wilson et al. (2017). Future studies should also analyze the results of the studies to determine which types of real-world connections are most relevant and beneficial for learners with special educational needs.

This study has several limitations. Future research should expand the time frame and the terminology related to special educational needs in the search strings to provide a broader overview of the field. In addition, since the review included only English-language sources, and terminology in special needs education varies significantly across countries, national research on real-world connections in special needs education was not covered.

References

- Bowman, J. A., McDonnell, J., Ryan, J. H., & Fudge-Coleman, O. (2019). Effective mathematics instruction for students with moderate and severe disabilities: A review of the literature. *Focus on Autism and Other Developmental Disabilities*, 34(4), 195–204.
- Browder, D. M., Spooner, F., Lo, Y., Saunders, A. F., Root, J. R., Ley Davis, L., & Brosh, C. R. (2017). Teaching students with moderate intellectual disability to solve word problems. *The Journal of Special Education*, 51(4), 222–235.
- Grünke, M., & Cavendish, W. (2016). Learning disabilities around the globe: making sense of the heterogeneity of the different viewpoints. *Learning Disabilities: A Contemporary Journal*, 14(1), 1–8.
- Julie, C. (2019). Practicing teachers and the development of mathematical modelling competencies through mathematical modelling as content. In C. Julie, L. Holtman, & C. R. Smith (Eds.), *Caught in the act: Reflections on continuing professional development of mathematics teachers in a collaborative partnership* (pp. 87–109). African Sun Media.
- Kaiser, G. (2017). The teaching and learning of mathematical modeling. In J. Cai (Ed.), *Compendium for research in mathematics education* (pp. 267–291). National Council of Teachers of Mathematics.

- Listiawati, N., Sabon, S. S., Siswantari, Subijanto, Wibowo, S., Zulkardi, & Riyanto, B. (2023). Analysis of implementing realistic mathematics education principles to enhance mathematics competence of slow learner students. *Journal on Mathematics Education*, 14(4), 683–700.
- Myers, J. A., Witzel, B. S., Powell, S. R., Li, H., Pigott, T. D., Xin, Y. P., & Hughes, E. M. (2022). A meta-analysis of mathematics word-problem solving interventions for elementary students who evidence mathematics difficulties. *Review of Educational Research*, 92(5), 695–742.
- Niss, M., & Blum, W. (2020). *The learning and teaching of mathematical modelling*. Routledge.
- Powell, S. R., Urrutia, V. Y., Berry, K. A., & Barnes, M. A. (2020). The word-problem solving and explanations of students experiencing mathematics difficulty: A comparison based on dual-language status. *Learning Disability Quarterly*, 45(1), 6–18.
- Schalock, R. L., Luckasson, R., & Tassé, M. J. (2021). An overview of intellectual disability: Definition, diagnosis, classification, and systems of supports (12th ed.). *American Journal on Intellectual and Developmental Disabilities*, 126(6), 439–442.
- Scott-Wilson, R., Wessels, D. C. J., Wessels, H. M., & Swart, E. (2017). The hidden benefits of mathematical modelling for Students with Disabilities. In G. A. Stillman, W. Blum, & G. Kaiser (Eds.), *Mathematical Modelling and Applications* (pp. 455–465). Springer International Publishing.
- Shin, M., Bryant, D. P., Powell, S. R., Jung, P.-G., Ok, M. W., & Hou, F. (2021). A meta-analysis of single-case research on word-problem instruction for students with learning disabilities. *Remedial and Special Education*, 42(6), 398–411.
- Spooner, F., Saunders, A., Root, J., & Brosh, C. (2017). Promoting access to common core mathematics for students with severe disabilities through mathematical problem solving. *Research and Practice for Persons with Severe Disabilities*, 42(3), 171–186.
- United Nations. (2007). *Convention on the rights of persons with disabilities*.
- Verschaffel, L., Greer, B., & de Corte, E. (2000). Making sense of word problems. *Contexts of learning*. Swets & Zeitlinger.
- Verschaffel, L., Schukajlow, S., Star, J., & van Dooren, W. (2020). Word problems in mathematics education: a survey. *ZDM*, 52(1), 1–16.
- World Health Organization. (2019). *International statistical classification of diseases and related health problems* (11th ed.).

MOVING BEYOND MOVES: REDEFINING MATHEMATICALLY RESPONSIVE TEACHING AS A HUMANISTIC STANCE

Gil Schwartz

Weizmann Institute of Science

Mathematically Responsive Teaching (MRT) is widely recognized as a cornerstone of high-quality instruction. However, its common definition is narrowly focused on specific teacher moves (e.g., eliciting student thinking). This theoretical paper redefines MRT to consider student experiences and needs alongside their mathematical ideas. Methodologically, it advocates moving beyond teaching moves to recognize a responsive stance toward students and emphasize authentic relationships with them, rooted in Buberian principles. The identification of an MRT stance is demonstrated through an analysis of a 10th-grade low-track statistics lesson.

BACKGROUND, RATIONALE, AND GOALS

Responsive teaching is everywhere and nowhere. It is often discussed as essential for student dialogic learning, but remains elusive, inconsistently defined, and backgrounded in research around high-quality mathematics teaching. Before unpacking this argument to address knowledge gaps and propose research goals, let us first consider how mathematically responsive teaching (MRT) is currently described:

- As “a characteristic of discourse that reflects the extent to which students’ mathematical ideas are present, valued, attended to, and taken up as the basis for instruction” (Bishop et al., 2022., p. 10).
- As “understanding students’ mathematical thinking and using it well in instruction” (Kim, 2019, p. 344).
- As “effectively respond[ing] to student ideas in ways that build disciplinary thinking” (Munson, 2019, p. 1).
- As “teaching in which teachers’ instructional decisions about what to pursue and how to pursue it are continually adjusted during instruction in response to children’s content-specific thinking” (Jacobs & Empson, 2016, p. 185).

Synthesizing these descriptions, a combined definition of MRT might be:

An instructional approach that foregrounds, in-the-moment, the substance of students’ mathematical ideas, attends to them as central to the learning process, and makes explicit connections to mathematical concepts and practices.

The descriptions above, while significantly advancing our understanding of MRT, have limitations in their unit of analysis and focus. They often lack detail on mathematical ideas, remaining general and detached from specific disciplinary knowledge. At the same time, they overemphasize predefined teacher moves, potentially overlooking broader, macro-level instructional goals. This raises questions about:

- What is “mathematical” in mathematically responsive teaching, and which elements of mathematical discourse do teachers attend to when building on student ideas? (micro-level questions)
- Are there alternative ways to conceptualize MRT beyond a narrow focus on specific teaching moves (e.g., probing, eliciting), and in-the-moment interactions? (macro-level questions)

My argument is that the current scholarly focus on analyzing MRT at a single level faces a dual limitation: it is both too broad and too narrow. It captures only a fragment of MRT's potential while dismissing instruction that falls outside its framework as non-responsive. While this observation and the two questions above suggest several important avenues for exploration, this paper specifically addresses the second question – the “too narrow” problem – by offering theoretical foundations to expand our understanding of MRT. To begin, I examine the omnipresence and elusiveness of MRT in the field.

Responsive teaching is everywhere. In Western countries, it is deeply embedded in policy documents, instructional frameworks (e.g., “the five practices”, Smith & Stein, 2018), and professional development programs, that emphasize the importance of attending to student ideas and grounding instruction in them. Even when not explicitly labelled as such, principles of MRT are evident across various educational approaches, including reform-based teaching, progressive pedagogy, and problem-based learning.

Notwithstanding, responsive teaching is nowhere. Searching the term in Google Scholar (in December 2024) yields illuminating results: within the top 30 results, 12 refer to culturally responsive mathematics teaching, and two to language-responsive mathematics teaching. This evidence highlights the notable paucity of research explicitly addressing MRT, but also sensitizes us, as a field, to the critical importance of attending to the cultural, relational, and social nature of learning, even when the focus is on the mathematical substance of student discourse. Consequently, and together with the limitations discussed above, an attempt to redefine MRT could be:

An instructional approach that foregrounds how teachers consider and act upon student mathematical thinking, perspectives, and experiences, when planning and teaching lessons.

Building on this definition of MRT as both seizing and actively creating opportunities to be receptive to students, in their whole beings, my overarching goal is to conceptualize MRT as a stance integrating teachers’ macro-decisions, moves, and micro-decisions, with attention to specific mathematical content. This paper focuses on the macro-goals lens by addressing a theoretical-methodological question: *How can we theorize and capture instances of MRT beyond predefined teaching moves?*

My motivation for addressing this issue stems from a phenomenon I noticed when analyzing videotaped mathematics lessons with researchers from different cultures. Oftentimes, lessons that some researchers viewed as highly responsive were criticized by others because they lacked certain moves, or their student talk seemed limited. This

paper seeks to address such discrepancies by broadening the theoretical framing of MRT. Moving beyond prescriptive rules about what teachers should or should not do in specific moments, I argue that MRT is a comprehensive stance that focuses on relationships with students around mathematical interactions. I suggest that MRT is a nuanced practice that requires weighing multiple factors, including the needs of individual students balanced with those of the group; fluency of procedures to support conceptual understanding; and considering, alongside curricular goals, relational goals such as fostering confidence, participation, or a sense of belonging. Thus, MRT is better understood as a stance, which is expressed differently in different situations, rather than as a fixed set of teacher moves. Methodologically, it means combining bottom-up analysis with a theoretical framing that centers relationships.

THEORETICAL FRAMING: A HUMANISTIC LENS ON MRT

In the literature on dialog and responsiveness, the primary sources of inspiration are often the works of Vygotsky, Bakhtin, and Freire, who emphasized the importance of social interactions and dialogic processes in teaching and learning (with the latter emphasizing education's transformative potential in fostering critical consciousness and empowering marginalized voices). However, to highlight the relational dimension of MRT, I seek to foreground the contributions of Martin Buber, a philosopher whose influence on educational thought, while perhaps indirect, is profound. Buber's approach, particularly his emphasis on authentic encounters *between* individuals, provides a rich philosophical foundation for understanding MRT as rooted in respect, mutuality, and genuine human connection. Buber (1958) distinguishes between two fundamental ways of relating to the world: I-Thou (you) and I-It.

- **I-Thou** is a relationship of mutual presence and dialogue, where both parties fully recognize and engage with each other as unique, whole beings. It emphasizes connection, empathy, and shared meaning, and involves deeply engaging with the experience of the other.
- **I-It** is a utilitarian relationship, where the other is treated as an object or means to an end, focusing on functions or roles rather than genuine interactions.

Importantly, Buber rejects a strict dualism between the I-Thou and I-It relationships. Instead, he is interested in their interplay and fluctuations. As noted by Morgan and Guilherme (2013), “the I-Thou relation will inevitably slip into an I-It relation, but the I-It relation has always the potential of becoming an I-Thou relation” (p. 5). They also state that “I-Thou experiences may be relatively few and far between” (ibid).

The concept of I-Thou relationships offers a rich lens for understanding MRT, as it emphasizes genuine interaction where each participant fully acknowledges the other's humanity. Considering the asymmetrical nature of teaching, MRT might be seen as an act of attunement, where the teacher actively takes the student's perspective, seeking to understand their thoughts, feelings, and struggles, as Buber notes:

“In dialogue, no matter whether spoken or silent . . . each participant really has in mind the other or others in their present and particular being and turns to them with the intention of establishing a living mutual relation between himself and them” (Buber, 1947, p. 18).

So, how might we envision mathematics instruction that is based on Buberian ideas? Three key principles emerge: First, recognizing the uniqueness of each student and respecting this uniqueness. Second, the importance of turning toward the Other, engaging with students as whole individuals. Third, ensuring that the engagement with the Other is authentic, fostering genuine interactions rather than superficial exchanges. While these principles focus on the teacher, they also emphasize the “between” — the relational space. The relational focus has implications for the identification of MRT, reinforcing the earlier critique of the axiomatic unit of analysis of “moves”. While traditionally, teaching moves and student responses constitute the basic unit of analysis for MRT studies (e.g., Bishop et al., 2022), Buber’s ideas call for identifying what is happening in the *sphere of between* the interlocutors, as Stengel (2003) suggests: “[pedagogical] action is not individual but relational. The unit of analysis is the relation, grounded always in Buber’s ‘I saying Thou’” (p. 196). This means, for example, stop placing undue emphasis on quantifying how many words students speak, but rather on the quality and authenticity of interactions. A similar point is made in Radford’s (2016) critique of efforts to promote student participation by reinforcing teacher/student dichotomies, which he argues are rooted in a capitalistic logic:

“The reform’s view of mathematics classroom knowledge production revolves around the idea of students’ participation. Although this is certainly a commendable idea, we see that students’ participation is usually understood against the backdrop of a *dichotomy* between teachers and students (p. 262)” [...] “any comment, question, or action from the teacher is perceived as influencing the students’ own ideas – a teacher’s deed that jeopardizes the students’ learning autonomy” (p. 263).

Building on these rich foundations, I want to focus, in this paper, on one element of MRT: teachers’ attunement to the experience of the Other, that is, to how their students perceive and experience classroom situations. To frame this, I draw on Wertsch’s (1984) concept of *situation definition*: “the way in which a setting or context is represented – that is, defined – by those who are operating in that setting [...] they often understand this context in such different ways that they are not really doing the same task” (pp. 8–9). Below, I examine instances where a teacher seemed to understand the context in ways that align with, or at least attempt to align with, his students’ perspectives, both in planning and enacting a lesson. I argue that these moments represent “I-Thou” instructional interactions, which embody a broader picture of MRT.

CAPTURING RESPONSIVE STANCE: AN ILLUSTRATION

This section presents a lesson analysis grounded in the theoretical framing outlined above. I first discuss the MRT decisions involved in planning and structuring the entire lesson – what I consider as beyond-the-moment MRT – and then analyze a 5-minute vignette from the lesson, presented verbatim with some abbreviations. The lesson and

the vignette were selected because they highlight a controversial moment for observers of the lesson, where a mathematical opportunity was seemingly missed by the teacher, potentially suggesting a moment of low responsiveness.

Beyond-the-moment MRT: structuring lessons around students' perspectives

MRT should be viewed not only as the act of seizing in-the-moment opportunities, but also as the intentional effort to summon those opportunities. This includes, but is not limited to, choosing relevant and rich problems, and selecting and sequencing student ideas to be able to build on them for reaching the lesson goal. Below is an example:

In a low-track 10th-grade statistics lesson (45 minutes; filmed in 2015 for the VIDEO-LM website), the teacher Ben, an experienced educator who is also the school's principal, introduced a problem from the matriculation exam to deepen students' understanding of averages, medians, and their properties. The question asks: "given four numbers (7, 10, 15, 16) and an additional number X between 10 and 15, find X if it is known that the average of the five numbers equals their median". After launching, the teacher monitored students' independent work for about 10 minutes. Next, he facilitated a class discussion where students presented, according to his selecting and sequencing, diverse strategies. According to him, Ben's teaching reflects a welcoming approach to informal mathematical reasoning in low-track classrooms, presented by Karsenty et al. (2007), which he previously adopted through their PD.

Ben's MRT decisions in structuring the lesson are evident in his careful selection of a problem. By choosing an exam-based problem whose solution reveals interesting mathematics, he addresses students' immediate goals while fostering deeper engagement with the concepts of average and median. Furthermore, the teacher sequences student solutions strategically, discussing first strategies that assist in understanding properties: The presentation of solutions begins with trial and error, a method accessible to most students and relatable as a starting point. It then progresses to a solution that uses properties of averages and medians without relying on variables, and finally introduces the construction of an equation as a formalized method—though not a mandatory or a desirable one.

In-the-moment MRT: Genuine interactions around mathematical perspectives

In this section, I present a vignette that follows the trial and error strategy, where Ben asks a student, Shiran, to explain her approach. To illustrate bottom-up categories inspired by the theoretical framing, I include a color-coded verbatim text:

Green: Recognizing students as authors of mathematical ideas; **Purple:** justifying ideas based on students' logic or collective benefit; **Turquoise:** accessible discussions of the mathematical ideas of the lesson – how average and median change; **Yellow:** honesty, transparency, and meta-talk about intentions; **Gray:** calculations.

1 Ben:	Shiran discovered something else. There's some kind of strategy here. I'm not sure if this is exactly what you meant, Shiran, but let's celebrate it because I want to say something about it. Shiran, do you want to explain what you meant?
2 Shiran:	$48=16+15+10+7$, and I divided it in the sum of these numbers, which is four. The result is exactly 12
3 Ben:	Look, if without adding the x, the average of these four numbers is 12... I'm not sure if that's exactly what Shiran meant, but we'll give her credit. Shiran, even if you didn't mean it, just act like you did, okay? Even though I say it. She says, if I, my average is 12, and this connects with what you, Ori, said earlier. I want you all to listen carefully because this is important now. There's a question like this coming up later in our question collection. If I have an average... I'll write this for you at the end, if I have a certain average, let's say my average in math this semester is 80. If I score 70 on the next test, what happens to the average?
4 Students:	It goes down.
5 Ben:	If I score 90, what happens to the average?
6 Students:	It goes up.
7 Ben:	Goes up. If I score 80, what happens to the average?
8 Students:	It stays the same.
9 Ben:	It stays the same. In other words, this is another way to solve this [...] Let's assume, in a different question, you weren't told... that x is between 10 and 15. You weren't told that. You're only given, uh... some numbers. You're asked: "Choose a number that will make the average higher". So, you calculate the average, Shiran took the numbers, calculated the average, and said, "The average is 12. What number do I need to add so the average doesn't change?"
10 Students:	12.
11 Ben:	12. That's already a mathematical method.
12 David:	Can I say why this isn't good?
15 Ben:	This is an excellent method... oh, do you want to explain why it's not good? Okay, you'll explain why it's not good in a moment. One second, I'll just write it down. It was said here that "if the average is 12, in order for the average to stay the same, we need to add 12" (writes on the board)—and that's our answer. Why isn't it good, David?
16, 18 David:	First, we were told to use five numbers. [...] and the median comes out to 12.5. If she does not explain it like this....
19 Ben:	The median isn't 12.5.
20 Student:	The median comes out to 12.5.
21 Ben:	Oh, nice. Nice.
22 David:	The thing is that the average and the median, they are equal.
23 Ben:	Okay, so look at what... look at what David is saying, which is something important for you to know. Here, after we found all the numbers and we know that it's 12, we know the median is 12. He says – here, even though the average is 12, what is the median here?
24 David:	12.5.
25 Ben:	How did you know it was 12.5, by the way?
26 David:	I checked.
27 Ben:	How did you check? Explain.
28 David:	I added them. What did I do? 10 plus 15, dividing by 2.
29 Ben:	The method is great [...] You balance out until you're left with two numbers. If there's just one number left, that's the median. Remember that...? Should we review this issue? [...] How does [your previous teacher] teach the median? She says: you arrange all the numbers in order, and then you start balancing—one from here, one from there, one from here, one from there. If one is left, then that's the median. In this case, says David, one from here, one from there, and we're left with two. If it's two, what do we do with them?
30 Students:	divide by two.
31 Ben:	You calculate the average. Got it? And indeed, the median here is 12.5. By the way, I don't know if you remember, but this is the reason why there is [...] not just the average but something else. Because we might see one factory where the average salary is 8,000 or 10,000 shekels and another factory where the average salary is 5,000 shekels. And we might say, "Wow, it's great for the workers in the 8,000-shekel factory and a shame for the workers in the 5,000-shekel factory." But then we discover that the salary in the supposedly "poorer" factory is actually much better. Why? What happens in the 8,000-shekel factory? There's one CEO or manager earning an extremely high salary. That's why they invented the median, you see? It's an example of how numbers can have the same average but a different median. Thank you, David, for that comment. Still, this method is a good one. I want you to remember this: If you've calculated a certain average, in order to maintain that same average, you need to add the same number. Now we turn to our third method, which Omer used...

This episode, when viewed through a deficit lens, might be dismissed as dominated by Initiate-Response-Evaluate patterns, excessive teacher talk, and an oversight of

David's idea (Line 12), missing its potential to discuss how the median changes. However, adopting the framing of responsive stance and I-Thou relationships, while recognizing the classroom's low-track context and the students' needs, provides insight into how MRT manifests in such environments. For instance, the utterances marked in purple highlight how the teacher validates students' perspectives, offers rationales grounded in their logic and experiences, provides intuitive examples, and addresses their concerns proactively ("I'll write this for you at the end", 3). The green utterances demonstrate Ben's revoicing which recognizes students as the authors of ideas. When David poses a potentially derailing question (12), Ben genuinely listens (21), skilfully reframes it to prevent confusion, and uses it to enhance students' procedural fluency (25, 29). Most critically, Ben dedicates significant time to discussing properties and their application to solving the problem, emphasizing that the second strategy is "a mathematical method" (11), not prioritizing equations as the sole focus.

While explicit eliciting or probing moves are limited, a la traditional MRT frameworks (e.g., Jacobs & Empson, 2016), this analysis demonstrates how Ben addresses students' mathematical perspectives, anticipates concerns, and fosters motivation, all while maintaining a focus on key lesson objectives. These moments, and the lesson's careful planning, illustrate what I-Thou educational relationships might look like.

CONCLUDING REMARKS

This paper sought to redefine MRT through a broader, humanistic lens, emphasizing the importance of recognizing students as whole individuals with unique experiences, needs, and perspectives. By shifting the focus beyond narrow interpretations of responsiveness, this approach highlights the value of fostering meaningful teacher-student interactions that honor students' lived realities and ways of thinking. Building on this foundation, the paper advocates for expanding our notion of best practices and adopting alternative units of analysis and methodologies.

This is only the beginning of a larger conversation about reimagining MRT through a humanistic framework while focusing on its subject-specific components. Future work will address also the micro-questions brought above by delving deeper into the mathematical intricacies of MRT and examining how they interrelate with broader pedagogical goals. Overall, expanding our conceptualization of MRT offers a more nuanced understanding of how it functions and manifests, especially in environments where it might otherwise go unnoticed.

Acknowledgment. I thank Ronnie Karsenty and Abraham Arcavi for granting permission to use the lesson transcript from the VIDEO-LM website, and Jason Cooper, Talli Nachlieli, and Avital Elbaum-Cohen for their inspiring workshop around this lesson at JCRME 12. Finally, I thank the teacher, Ben, whose teaching provided me with unforgettable lessons.

References

- Bishop, J. P., Hardison, H. L., & Przybyla-Kuchek, J. (2022). Responsiveness to students' mathematical thinking in middle-grades classrooms. *Journal for Research in Mathematics Education*, 53(1), 10–40. <https://doi.org/10.5951/jresmetheduc-2020-0188>
- Buber, M. (1947). *Between man and man*. The Macmillan Co.
- Buber, M. (1958). *I and Thou*. Scribner.
- Jacobs, V. R., & Empson, S. B. (2016). Responding to children's mathematical thinking in the moment: An emerging framework of teaching moves. *ZDM – Mathematics Education*, 48, 185–197. <https://doi.org/10.1007/s11858-015-0717-0>
- Karsenty, R., Arcavi, A., & Hadas, N. (2007). Exploring informal mathematical products of low achievers at the secondary school level. *The Journal of Mathematical Behavior*, 26(2), 156–177. <https://doi.org/10.1016/j.jmathb.2007.05.003>
- Kim, H. J. (2019). Teacher learning opportunities provided by implementing formative assessment lessons: Becoming responsive to student mathematical thinking. *International Journal of Science and Mathematics Education*, 17(2), 341–363. <https://doi.org/10.1007/s10763-017-9866-7>
- Morgan, W. J., & Guilherme, A. (2013). *Buber and education: Dialogue as conflict resolution*. Routledge.
- Munson, J. (2019). After eliciting: Variation in elementary mathematics teachers' discursive pathways during collaborative problem solving. *The Journal of Mathematical Behavior*, 56, 100736. <https://doi.org/10.1016/j.jmathb.2019.100736>
- Radford, L. (2016). On alienation in the mathematics classroom. *International Journal of Educational Research*, 79, 258–266. <https://doi.org/10.1016/j.ijer.2016.04.001>
- Smith, M., & Stein, M. K. (2018). *5 Practices for orchestrating productive mathematics discussion*. National Council of Teachers of Mathematics.
- Stengel, B. S. (2003). "As If We Were Called": Responding to (Pedagogical) Responsibility. *Philosophy of Education Archive*, 195–203. <https://doi.org/10.47925/2003.195>
- Wertsch, J. V. (1984). The zone of proximal development: some conceptual issues. *New Directions for Child and Adolescent Development*, 23, 7–18. <https://doi.org/10.1002/cd.23219842303>

PRE-SERVICE PRIMARY TEACHERS' DIAGNOSTIC JUDGEMENTS: EFFECTS OF TEACHER DISPOSITIONS AND STUDENT CHARACTERISTICS

Tjorven Lea Seibold, Lars Meyer-Jenßen, Christin Laschke, Bettina Roesken-Winter

University of Münster, Leibniz-Institute for Science and Mathematics Education,
Health and Medical University Erfurt, University of Münster

Diagnostic judgments may be influenced by teacher dispositions and student characteristics. To this end, a survey was conducted on 181 pre-service primary teachers regarding their diagnostic judgments in the domain of quantities and measurement, investigating the effects of math anxiety and mathematical content knowledge (teacher dispositions) as well as gender, migratory background and socioeconomic status (student characteristics) using path analysis. The findings revealed effects of math anxiety, mathematical content knowledge and socioeconomic status, but no correlations concerning gender and migratory background. This highlights the importance of awareness to potential bias about student characteristics and of imparting diagnostic knowledge as part of teacher education.

INTRODUCTION

The ability to make accurate diagnostic judgments is one of the key qualifications of a teacher (e.g. Leuders et al., 2018). For instance, diagnostic judgements can build the foundation of planning and implementing adaptive learning environments to foster students' understanding (Hardy et al., 2019). Furthermore, a positive correlation between the diagnostic competence of a teacher and students' achievement underscores the relevance of these judgements (Südkamp et al., 2012). However, studies have identified challenges teachers encounter in diagnosing students' learning, with personal dispositions and student characteristics potentially influencing their judgements (Urhahne & Wijnia, 2021). The present study aims to examine the extent to which these effects of personal dispositions and student characteristics on diagnostic judgments are already evident among pre-service primary teachers.

THEORETICAL FRAMEWORK

According to the DiaCoM framework, diagnostic judgements can be understood as “inferences of a teacher about students (e.g., ability) or materials (e.g., task difficulty) based on the information explicitly or implicitly given in a diagnostic situation” (Loibl et al., 2020, p. 2-3). Diagnostic judgements can be evaluated based on their accuracy, which can for example be measured by the extent to which teachers identify missing understanding and potential in a student's solution (Südkamp et al., 2012). In their

framework Loibl et al. (2020) put the aspects of perceiving, interpreting and decision-making, at the core of diagnostic judgments. The resulting product of this process they label as diagnostic behavior, thereby referring to the observable outcome of the diagnostic process. The framework delineates various characteristics that have the potential to influence the diagnostic process therefore also diagnostic judgements. These characteristics can be classified either as states or traits of the teacher, in the sense of teacher dispositions, or as arising from the context in which the diagnostical judgement unfolds, such as cues that suggests certain personal characteristics of the students (Loibl et al., 2020).

Concerning teacher dispositions, Philip (2018) has demonstrated that teachers primarily relied on their mathematical content knowledge (MCK) for diagnostic purposes. Math anxiety has been shown to have a negative effect on MCK, which can consequently also adversely affect their diagnostic judgements (Jenßen et al, 2019). With regard to student characteristics and their influence on the diagnosis of teachers, the research findings are mixed with regard to the characteristics of gender, migratory background and socioeconomic status (SES): Some studies demonstrate a correlation between these characteristics and teachers' diagnostic judgements, while other report no such effects (Urhahne & Wijnia 2021). With regard to gender, the studies generally indicate only minor or no effects at all on the diagnosis of teachers (Südkamp et al., 2012). Concerning the characteristic of migratory background, Glock (2016) demonstrated that in the domain of languages, low-performing students with a migratory background received lower ratings compared to students without a migratory background, yielding similar results. However, such different assessments did not occur in mathematics. Additionally, expectations of the performance of students with a migratory background are frequently lower than for other students (Costa et al., 2022). Regarding socio-economic status, Kaiser et al. (2015) found no correlation between students' SES and teachers' diagnostic judgements. However, Ready & Wright (2011) reported that teachers often hold lower expectations for student with low SES.

RESEARCH QUESTIONS

The research findings on the role of student characteristics and teacher dispositions for teachers' diagnostic judgments have thus far been limited to investigating in-service teachers. However, as shown by Costa et al. (2022), these effects may also extend for pre-service teachers. Consequently, we investigated the effects of both student characteristics (here ascribed characteristics to fictional students) and teacher dispositions on pre-service teachers' diagnostic judgments in the area of quantities and measurement. The following research question will be pursued:

RQ: What effect do ascribed characteristics of students (i.e., gender, migratory background, SES) and teacher dispositions (i.e., MCK, math anxiety) have on

diagnostic judgements of pre-service primary teachers regarding student's solutions to tasks in quantities and measurement?

H1: Pre-service teachers' MCK shows a positive effect on their diagnostic judgements (Philip, 2018).

H2: Pre-service teachers' math anxiety shows an indirect negative effect on their diagnostic judgements, mediated through their MCK (Jenßen et al., 2019).

Due to the inconclusive nature of the research findings regarding the effects of student characteristics (Urhahne & Wijnia 2021), no hypotheses are formulated concerning these variables.

METHODOLOGY

Participants and context

In the present study, 181 pre-service primary teachers from a university in a metropolitan region participated, of whom 82% were in their fifth semester, 3% in their third or fourth semester, and 14% in their sixth or higher semester of the Bachelor's degree course. While the majority of the pre-service teachers had not yet attended courses focusing on diagnostic aspects of mathematics teaching in depth, they had already been involved in this topic in other contexts or had carried out initial analyses of student solutions. The majority of the participants identified as female (77% female, 22% male, 1% diverse). The survey was administered online and was incorporated in the compulsory coursework for a lecture, and the pre-service primary teachers had the option of considering their contribution for scientific purposes.

Instrument

The pre-service primary school teachers were surveyed using a vignette-based test, focusing on the topic "quantities and measurements". Math anxiety was assessed using four items (Jenßen et al., 2021). To situate pre-service teachers' diagnostic judgements, they were asked to envision themselves as teachers in a new class, where quantities and measurements would be the next topic. To assess prior knowledge of the class, 20 tasks were assigned. In order to assess pre-service teachers MCK in the domain of quantities and measurement, they first solved the tasks themselves. One task, for instance, asked whether a specific ruler could be used to draw a line of 4 *cm* length. The ruler in questions featured uneven distances between the numbers, rendering the centimeter unit unusable for measurement. In the following, the pre-service primary school teachers were presented a student solution to these tasks, and were asked to diagnose the student's level of competence in an open format. This served to operationalize the processes of perceiving and interpreting in the context of diagnostic judgements. Regarding the ruler task, the presented solution was as follows:

"Yes, that's easy. I draw a line from the beginning to four."

The student did not yet understand that the centimeter unit cannot vary in length, meaning the numbers on the ruler representing this unit must appear in equidistant lengths.

All solutions to the 20 tasks were structured similarly to the example above, indicating the student has not yet fully grasped the subject matter in question. Finally, the pre-service teachers had to choose one out of four prompts to react to the student's solution. Out these four prompts, only one was accurate from a pedagogical content knowledge (PCK) perspective, fostering students' understanding. The others contained PCK aspects not relevant for the situation at hand, were merely pedagogically oriented or senseless from a mathematical viewpoint. The selection of the accurate PCK-prompt was used to assess the pre-service teachers' decision making as part of their diagnostic judgement.

To investigate the effects of student characteristics on pre-service teachers' diagnosis, the preservice primary teachers were randomly assigned to different student characteristics. However, the student's solutions were kept the same regardless of the student assigned.

The ascribed student characteristics included information on gender (female/male), migratory background (German/Arabic), and SES (low/high), and were assigned by name, written information (e.g. parents' occupation, living situation, hobbies), and a drawing of the respective student. The combination of characteristics resulted in eight different students (the four variants for the female gender shown in Figure 1).

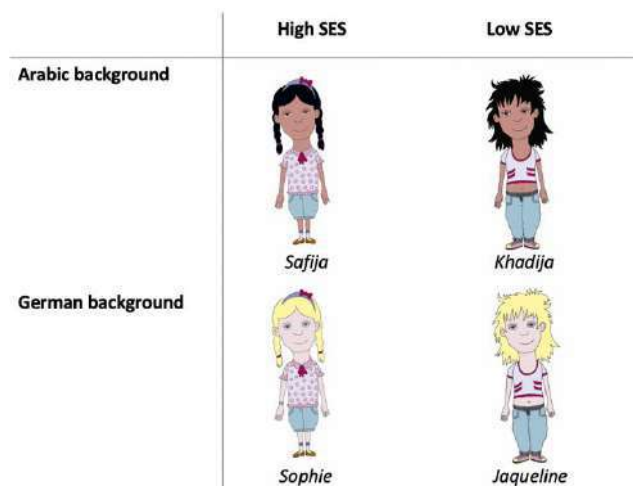


Fig. 1: Assigned Female students

Data analysis

The open-ended responses to diagnosing the solutions with respect to the student's level of competence were evaluated using qualitative content analysis, with coding carried out deductively. The category *missing understanding identified* was assigned when pre-service teachers revealed that the student solution showed a lack of understanding. The category was not assigned when the lack of understanding described did not match the student answer. The category *potential identified* was assigned to pre-service teachers who described the student's already existing skills and knowledge. For instance, in the ruler task, it could be stated that the student is probably able to draw a 4 cm long line with a functioning ruler. Once again, the category was not assigned if the potential described did not match the student's answer. The responses of 40 pre-service teachers (representing over 20% of the sample) were also

coded independently by another researcher to ascertain interrater reliability ($\kappa = .87$).

The coding process was executed for each of the 20 student solutions, with a score of 1 assigned to each category and a score of 0 assigned to categories that were not met. The maximum attainable score for each category was 20. The effects of the characteristics addressed in the vignettes (SES, gender, migratory background) and the pre-service teachers' dispositions (math anxiety, MCK) on their diagnostic judgements were then transferred into a model using path analysis (Stage et al, 2004). This modeling was carried out using the MPlus software (Muthen & Muthen, 1998-2017).

RESULTS

Overall, the calculated model (Fig. 2) showed a good fit to the data ($\chi^2(6) = 9.269$, $p = .41$, RMSEA = .01 [.00/.09], CFI = .99, SRMR = .03). All effects shown in the model are significant.

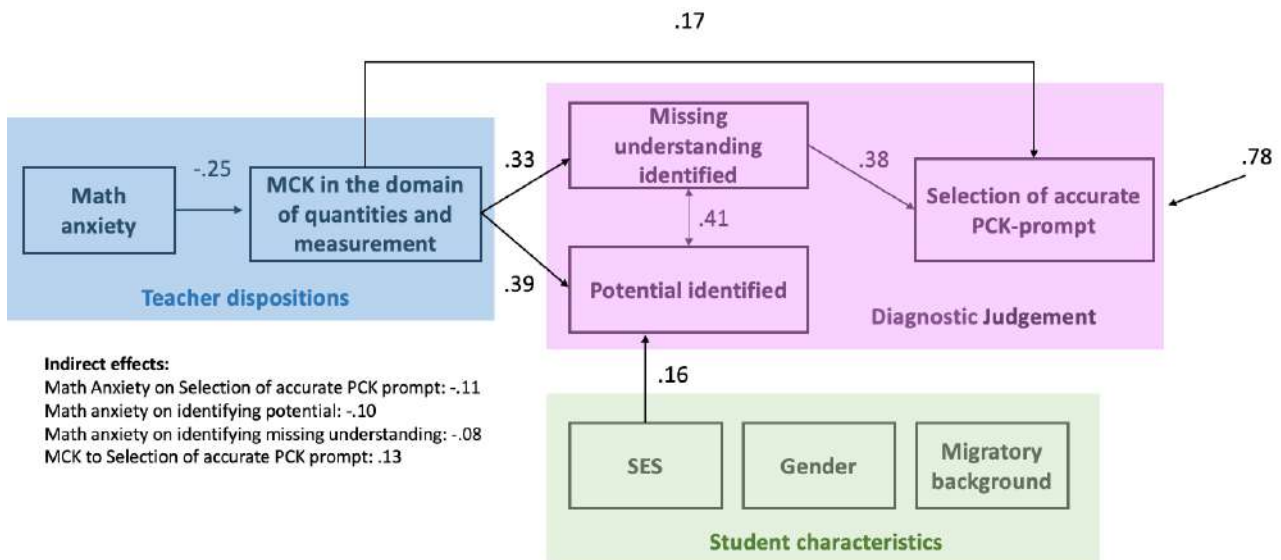


Fig. 2: Path model regarding effects of student and teacher characteristics on diagnostic judgements in the domain of quantities and measurement

Regarding the different aspects of diagnostic judgements, a positive correlation between *missing understanding identified* and *potential identified* was shown ($\beta_{ind} = 0.41$, $p < .001$). Furthermore, there was a significant effect of *missing understanding identified* on the selection of the accurate PCK-prompt ($\beta_{ind} = 0.38$, $p < .001$). Significant effects were further identified with regard to pre-service teachers' MCK, influencing their diagnostic judgements. On the one hand, positive correlations were observed between MCK and *missing understanding identified* ($\beta_{ind} = 0.33$, $p < .001$) and *potential identified* ($\beta_{ind} = 0.39$, $p < .001$) in the students' solutions. On the other hand, the pre-service teachers' MCK exhibited a significant effect on the selection of the accurate PCK-prompt ($\beta_{ind} = 0.17$, $p < .001$). As anticipated, a negative correlation between mathematics anxiety and pre-service teachers' MCK was identified ($\beta_{ind} = -$

0.25, $p < .001$). Consequently, indirect negative effects of mathematics anxiety were observed on *missing understanding identified* ($\beta_{ind} = -0.80, p < .001$) and *potential* ($\beta_{ind} = -0.10, p < .001$) as well as on the selection of the accurate PCK-prompt ($\beta_{ind} = -0.11, p < .001$). Concerning the student characteristics, potential in the student solutions was more likely to be identified for students ascribed a high socioeconomic status than for students ascribed a low SES ($\beta_{ind} = 0.16, p < .001$). No correlation was found between the students' SES and *missing understanding identified*. Regarding gender or migratory background of the students no significant effects on diagnostic judgements of the pre-service primary school teachers were observed.

To further illustrate the disparities revealed by pre-service teachers' open answers, we examine responses provided for the ruler task outlined in the methods section. The same student solution, ascribed a low SES (Jacqueline) and a high SES (Sophie), yielded divergent replies:

Jacqueline has not yet understood that the units always have the same distance between them. The only thing that counts for her is the number at the end that has to be reached.

Sophie has understood that the 4 on the ruler normally represents the unit 4 cm. However, she is not yet aware of how the millimeter displays (lines) and the centimeters relate to each other and therefore the display on this ruler has been falsified or does not reflect a distance of 4 cm.

The utterances in question align with the category *missing understanding identified*, as they accurately reflect the student's lack of understanding regarding the constant length of the unit centimeter. However, the response to Sophie's solution addresses her already existing knowledge, and underscores the potential that can be derived from her solution. Specifically, the pre-service teacher stated that Sophie is aware of the fact that the numeral 4 on a ruler typically signifies 4 centimeters.

A parallel discrepancy emerged in the responses to the same student solution, attributed either to Hamza (low SES) or Omar (high SES):

Hamza has not yet understood that the same distance must be used everywhere.

Omar understands that a 4 on the ruler means that it is 4 cm from the start of the ruler to the end. However, he has not yet understood or has simply overlooked the fact that the distances between the ruler sections must be the same so that the 4 cm are really 4 cm.

In this case, both answers were assigned to *missing understanding identified*. However, only in the solution ascribed to Omar, the pre-service teacher *identified potential*. The error made by Omar was attributed to carelessness in reading the task.

Furthermore, responses classified as *missing understanding identified*, but not *potential identified* frequently lacked elaboration on the missing understanding. Instead, characterizing the student's performances as "low", "poor" or at the "beginner's level" dominated. Additionally, pre-service teachers maintained a general level when judging a solution attributed to a low-level SES:

Jaqueline has not yet fully understood the topic and does not yet know what to do with units of measurement.

DISCUSSION AND CONCLUSION

The results obtained in the present study provided evidence to support both hypotheses 1 and 2, which were derived from the context of in-service teachers: Pre-service teachers' MCK exhibited a positive effect on their diagnostic judgements. In contrast, math anxiety manifested in an indirect negative effect on pre-service teachers' diagnostic judgements, mediated through their MCK. Turning to the student characteristics, it was observed that students with high SES backgrounds were more likely to be ascribed with potential in their answers. This phenomenon manifests in a more positive perception and interpretation of the student's solution and often results in more detailed responses, as pre-service teachers also take prior knowledge into account. Thus, whether potential could be seen in a student answer was not dependent on the student utterances per se, but was influenced by an ascribed student background. Also, answers that only identify missing understanding are often formulated in a general way. Notably, SES is the only student characteristic that shows an effect on one part of the pre-service primary teachers' diagnostic judgments. This may possibly be due to the context of the study, as pre-service primary teacher training in a large city was surveyed. These pre-service teachers may have been more confronted with social diversity in their environment. Furthermore, the social discourse in recent years may have led to a heightened awareness among pre-service teachers concerning bias towards certain student characteristics. This is also one of the limitations of the study: the pre-service teachers were not sampled at random, but were a convenience sampling, which could have influenced some effects as described above. Overall, the results indicate that pre-service teachers should be sensitized in regards to potential bias concerning student characteristics in the context of diagnostic judgements. In addition, an in-depth discussion of mathematical content knowledge during teacher training can also contribute to better diagnostic judgements.

References

- Costa, S., Pirchio, S., Glock, S., 2022. Teachers' and preservice teachers' implicit attitudes toward ethnic minority students and implicit expectations of their academic performance. *International Journal of Intercultural Relations* 89, 56–62. <https://doi.org/10.1016/j.ijintrel.2022.05.006>.
- Glock, S. (2016). Does ethnicity matter? The impact of stereotypical expectations on in-service teachers' judgments of students. *Social Psychology of Education*, 19 (3), 493–509. <https://doi.org/10.1007/s11218-016-9349-7>.
- Hardy, I., Decristan, J., Klieme, E. (2019). Adaptive teaching in research on learning and instruction. *Journal for educational research online*, 11 (2), 169-191. <https://doi.org/10.25656/01:18004>.

- Jenßen, L., Thiel, O., Dunekacke, S., & Blömeke, S. (2019). Mathematikangst bei angehenden frühpädagogischen Fachkräften: Bedeutsam für professionelles Wissen und Wahrnehmung von mathematischen Inhalten im Kita-Alltag? [Early childhood teacher students' mathematics anxiety: Relevant for Professional knowledge and perception of mathematics in Preschool?] *Journal für Mathematik- Didaktik*, 41, 301–327. <https://doi.org/10.1007/s13138-019-00151-1>
- Jenßen, L., Eilerts, K., & Roesken-Winter, B. (2021). Differentielle Effekte sozialer Unterstützung auf das Erleben von Angst und Freude in Mathematik von Studierenden des Grundschullehramts [Differential effects of social support on the experience of anxiety and joy in mathematics of pre-service primary teachers]. In J. Stiller, C. Laschke & L. Goecke (Eds.), *Berlin-Brandenburger Beiträge zur Bildungsforschung 2022* (pp. 17-40). Humboldt-Universität zu Berlin. <https://edoc.hu-berlin.de/handle/18452/25292>.
- Kaiser, J., Möller, J., Helm, F., & Kunter, M. (2015). Das Schülerinventar: Welche Schülermerkmale die Leistungsurteile von Lehrkräften beeinflussen [The student inventory: How student characteristics bias teacher judgments]. *Zeitschrift für Erziehungswissenschaft*, 18(2), 279–302. <https://doi.org/10.1007/s11618-015-0619-5>.
- Leuders, T., Dörfler, T., Leuders, J., & Philipp, K. (2018). Diagnostic competence of mathematics teachers: Unpacking a complex construct. In T. Leuders, K. Philipp, & J. Leuders (Eds.), *Diagnostic competence of mathematics teachers* (Vol. 11, pp. 3–31). Springer International Publishing. https://doi.org/10.1007/978-3-319-66327-2_1.
- Loibl, K., Leuders, T., & Dörfler, T. (2020). A framework for explaining teachers' diagnostic judgements by cognitive modeling (DiaCoM). *Teaching and Teacher Education*, 91, 103059. <https://doi.org/10.1016/j.tate.2020.103059>.
- Muthen, L. K., & Muthen, B. O. (1998-2017). Mplus user's guide (8th ed.). Muthen & Muthen.
- Philipp, K. (2018). Diagnostic Competences of mathematics teachers with a view to processes and knowledge resources. In T. Leuders, K. Philipp & J. Leuders (Hrsg.), *Diagnostic competence of mathematics teachers* (S. 109–127). Cham: Springer.
- Ready, D. D., & Wright, D. L. (2011). Accuracy and inaccuracy in teachers' perceptions of young children's cognitive abilities: The role of child background and classroom context. *American Educational Research Journal*, 48(2), 335–360. <https://doi.org/10.3102/0002831210374874>.
- Stage, F. K., Carter, H. C., & Nora, A. (2004). Path analysis: An introduction and analysis of a decade of research. *Journal of Educational Research*, 98(1), 5–13. <https://doi.org/10.3200/JOER.98.1.5-13>.
- Südkamp, A., Kaiser, J., & Möller, J. (2012). Accuracy of teachers' judgments of students' academic achievement: A meta-analysis. *Journal of Educational Psychology*, 104 (3), 743–762. <https://doi.org/10.1037/a0027627>.
- Urhahne, D., & Wijnia, L. (2021). A review on the accuracy of teacher judgments. *Educational Research Review* 32, Article 100374. <https://doi.org/10.1016/j.edurev.2020.100374>.

LOCATING NUMBERS ON THE NUMBER LINE: AN EYE-TRACKING STUDY WITH FIRST GRADERS WITH AND WITHOUT RISK OF DEVELOPING MATHEMATICAL DIFFICULTIES

Anna Lisa Simon¹, Laura Geimer¹, Lukas Baumanns², Demetra Pitta-Pantazi³,
Constantinos Christou³, Achim J. Lilienthal⁴, & Maike Schindler¹

¹University of Cologne, ²TU Dortmund, ³University of Cyprus,

⁴TU Munich & Örebro University

The number line (NL) is an important tool in mathematics education. Previous research has found that fifth graders with mathematical difficulties show difficulties in locating numbers on marked NLs, which are common in mathematics education. The aim of this eye-tracking study is to investigate if these difficulties are already evident among students at the beginning of primary school. We investigated if first graders at risk of developing mathematical difficulties (RMD) (n=23) and first graders not at risk of developing such difficulties (n=100) differ in locating numbers on a marked NL from 0–20 – in success rates, response times, and strategies. We found that RMD students made more errors. While we did not find differences in response times, students' strategy use differed in that RMD students used less efficient strategies.

INTRODUCTION

The number line (NL) is an important mathematical tool, especially in primary school. Previous research has shown that students' performance in empty NL tasks correlates with their overall mathematical performance (Friso-van den Bos et al., 2015), making NL tasks interesting for research on mathematical difficulties. Students with mathematical difficulties (MD) tend to have difficulties in locating numbers on empty NLs (e.g., van't Noordende et al., 2016). For marked NLs, which are commonly used in mathematics education, students with MD at the beginning of secondary school were found to show difficulties in strategy use compared to students without MD (e.g., Simon et al., 2023). For students at the beginning of primary school, the NL is a complex matter to understand, especially for students who have difficulties in learning mathematics. They need to understand representations of numbers on an NL and how to use NLs. To date, little is known about how students handle marked NLs, which are introduced and used early on in mathematics education. A method that has proven to be useful in gaining insights into students' use of NLs is eye tracking (ET) (e.g., van't Noordende et al., 2016). In this paper, we use ET to investigate if the difficulties found among fifth graders with MD in strategy use for marked NLs are already evident in first-grade students. We study how first graders at risk of developing mathematical difficulties (RMD) and first graders not at risk (non-RMD) handle marked NL tasks.

THEORETICAL BACKGROUND

Number line

The NL is an important tool in mathematics education, particularly in primary school (Teppo & van den Heuvel-Panhuizen, 2014). By representing the basic concept of the number sequence, the NL supports the development of number sense and promotes a deeper understanding of ordinal numbers. Students can use the NL to explore the number order and relationships between numbers (Teppo & van den Heuvel-Panhuizen, 2014). There are various types of NLs, for instance, those with only labelled starting and endpoints (called empty or proportional NLs) and others that include additional hatch marks and labelled reference points, for example, fives and tens (depending on the number ranges), which we refer to as *marked* NLs. The use of reference points can have a positive effect on NL performance (Peeters et al., 2017).

Research has shown that students' performance in empty NL tasks correlates with their overall mathematical performance. It is assumed that students' overall mathematical performance and their performance in locating numbers on a NL influence each other (Friso-van den Bos et al., 2015), highlighting students' NL skills for their overall mathematical performance. Yet, the NL is a complex matter and often difficult to understand, especially for MD students (e.g., Bull et al., 2021).

(Risk of developing) Mathematical difficulties

Mathematical difficulties (MD) are understood as difficulties in basic arithmetic skills, for example, in understanding numbers and operations (Moser Opitz et al., 2017). Students with MD tend to have difficulties in empty NL tasks: they show lower accuracy in locating numbers and a less adequate use of strategies (e.g., van't Noordende et al., 2016). Already for six-year-old children, difficulties with empty NL tasks can hint at later MD (e.g., Bull et al., 2021).

MD become apparent during primary school and can lead to difficulties beyond school years, in every day and professional life (e.g., Geary, 2011). Since a reliable diagnosis of MD is usually possible from the second grade onwards, children who show difficulties in learning mathematics before this are considered to be at risk of developing mathematical difficulties (RMD). To address students' difficulties early on and support them appropriately, it is important to identify their strengths and difficulties. For marked NLs, little is known about how young students use these NLs and how they benefit from using this tool.

Eye tracking

To perceive objects of attention sharply, the human eye needs to move. The method of ET records these eye movements (Holmqvist et al., 2011) and allows the analysis of a person's visual attention (Carrasco, 2011). The visual attention of a person allows insights into how a person solves mathematical tasks and the analysis of ET data has been shown to be valuable for investigating students' strategies when working on

empty NL tasks (e.g., van't Noordende et al., 2016). Previous ET studies investigating fifth graders' NL strategies for locating numbers on marked NLs from 0–100 and 0–1000 showed that students with MD tended to use more counting strategies than direct strategies (Simon & Schindler, 2022) and used reference points differently (Simon et al., 2023) compared to students without MD.

To the best of our knowledge, there have been no ET studies yet that investigated how RMD first graders solve tasks on a marked NL from 0–20. The aim of this study is to investigate if first graders with and without RMD differ in locating numbers on a marked NL. We ask the research question: *Do RMD first graders and non-RMD first graders differ in locating numbers on a marked NL, in particular in (1) success rates, (2) response times, and (3) strategies used as inferred from eye movements?*

THIS STUDY

Participants, procedure, and setting

The study took place at three primary schools in Germany with first grades. All students participated in a standardized mathematics test, a short version of the ZAREKI-K (von Aster et al., 2009; adapted short version by Walter, 2020), to assess students' mathematical performance and to identify RMD students. For the analyses of NL tasks in this paper, 23 RMD students and 77 non-RMD students are considered. We used position-to-number-tasks (Fig. 1) on a marked NL from 0–20. Students were shown a red cross on the NL and asked for the corresponding number. In total, students worked on ten tasks (five tasks with single-digit numbers and two-digit numbers each) on the NL 0–20. We used the numbers 9, 15, 2, 12, 10, 3, 5, 14, 7, and 19. Tasks were presented to the students in the same randomized order as described above.

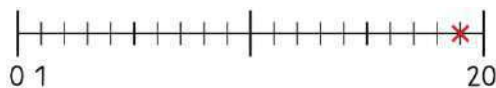


Figure 1: Example of a NL task (number 19)

The students worked individually in a quiet room at their school. Before the first task, they completed a warm-up task to ensure that they had understood the task correctly. The students answered by saying out loud the number they thought was indicated by the red cross. They did not receive any feedback on their answers. The students' eye movements were recorded by the remote eye tracker Tobii Pro X3-120 (binocular, infrared, 120 Hz). The students were seated approx. 60 cm away from a 24" monitor on which the tasks were presented. The ET data had an average accuracy of 0.96°.

Qualitative analysis of eye-tracking data

We used raw gaze-overlaid videos provided by Tobii Pro Lab software to analyze students' NL strategies. NL strategies were coded using a category system developed by Simon and Schindler (2022), which was (inductively) adapted (Fig. 2): In the study by Simon and Schindler, the starting point and endpoint of the NL were highlighted and labelled, and the midpoint was also highlighted (longer, thicker mark). In the study

presented here, the marks for the numbers 5 and 15 were also highlighted (thicker marks) (Fig. 1). In case students used these highlighted, but unlabelled reference points ('5', midpoint, i.e., '10', and '15'), this was coded as 'use of near(est) reference point'. Near(est) reference points can be both the reference point to the left and to the right of the target number (e.g., '5' and '10' for target number 7). We further distinguished whether, starting from the reference point, the students counted on (strategy 4) or located numbers directly (strategy 3) (see Fig. 2). To visualize strategies in this paper, we use gaze plots, although we used gaze-overlaid videos in our analyses. For all ET videos, NL strategies were coded by the second author. The first author coded 25 % of the ET videos independently. The interrater agreement was almost perfect ($\kappa = 0.91$).

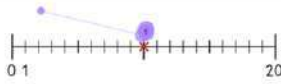

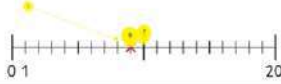


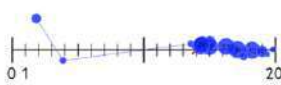
Strategy number	Strategy name	Example
Strategy 1	Direct locating	
Strategy 2	Starting point use and counting	
Strategy 3	Use of near(est) reference point and direct locating	
Strategy 4	Use of near(est) reference point and counting	
Strategy 5	Endpoint use and direct locating	
Strategy 6	Endpoint use and counting	

Figure 2: Adapted category system with examples of students' gazes from this study

Quantitative analysis

We analyzed success rates (proportion of correctly answered tasks). In our study, errors can mean either an incorrect answer or no answer by the student, for example, the student wanted to skip the task. For analyzing group differences of success rates, we used Mann-Whitney U test as non-parametric test because of non-normally distributed data (Shapiro-Wilks $p < .05$). We calculated effect sizes as Pearson's correlation coefficient r . For the analyses of response times and strategies, only tasks in which students' results did not deviate by more than 2 from the correct result (e.g., 18 or 19 instead of 17) were considered: When students, for example, miscounted by 1 (or 2, which can occur when counting forwards from the starting point of the NL for numbers towards the end of the NL) these tasks were considered, while tasks with more divergent results were not included in order to avoid analyzing strategies when students guessed the answer. The same tasks were considered for the analyses of response times and strategies. For the analyses of response times, we used Mann-Whitney U test. For the analyses of strategies, we used chi-square tests. We calculated effect sizes using Cramér's V . Bonferroni-Holm adjusted p -values are reported.

RESULTS

Success rates. The comparison of the success rates (Fig. 3) of all tasks together ($U = 267.50$, $Z = -5.22$, $p < .001$, $r = .52$) showed that non-RMD students answered significantly more tasks correctly than RMD students. The same was true if one looked separately into single-digit numbers ($U = 412.50$, $Z = -4.29$, $p < .001$, $r = .43$) and two-digit numbers ($U = 257.00$, $Z = -5.44$, $p < .001$, $r = .54$).

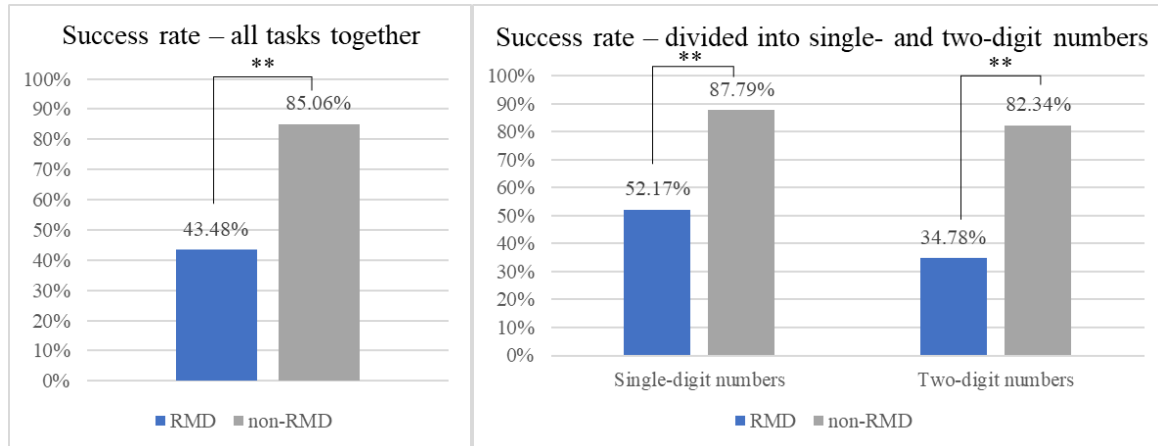


Figure 3: Students' success rates for all tasks together (left) and for tasks with single- and two-digit numbers separately (right) (** $p < .001$)

Response times. The comparison of the response times of all tasks together ($U = 692.00$, $Z = -.28$, $p = .780$) showed no significant differences between students with and without RMD. Neither did the comparison of the response times of the single-digit number tasks ($U = 699.00$, $Z = -.21$, $p = .831$), nor of the two-digit number tasks ($U = 514.50$, $Z = -.89$, $p = .373$).

Strategies. Chi-square test for all tasks together (Fig. 4) showed significant differences in the use of strategies between students with and without RMD: $\chi^2(5) = 30.39$, $p < .001$, $V = .19$.

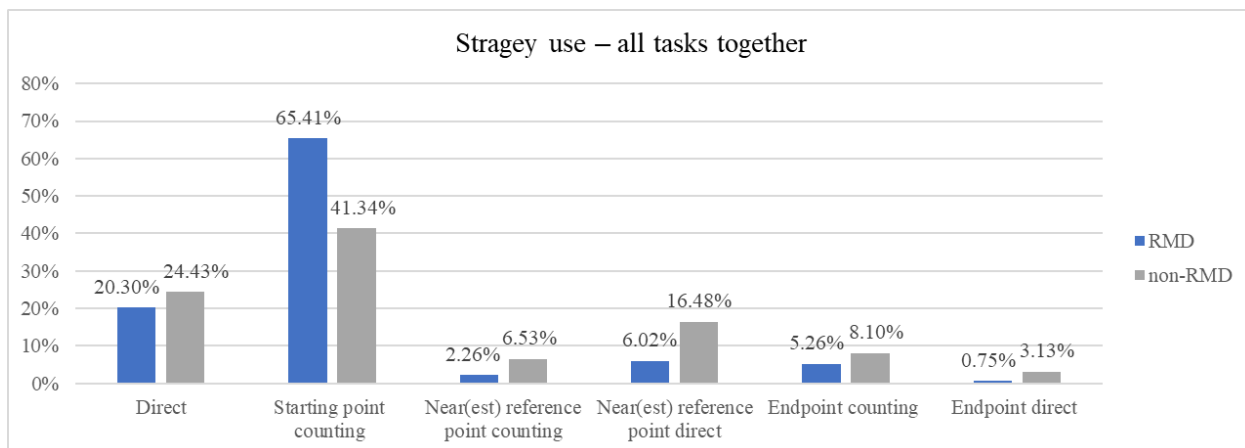


Figure 4: Students' strategy use for all tasks together

For single-digit numbers (Fig. 5), chi-square test showed significant differences in the use of strategies between students with and without RMD: $\chi^2(3) = 9.49$, $p = .023$,

$V = .15$. Cell tests showed that RMD students used strategy 2 ‘starting point use and counting’ significantly more often ($\chi^2(1) = 5.93, p = .030, V = .12$) and strategy 4 ‘use of near(est) reference point and counting’ significantly less often ($\chi^2(1) = 6.44, p = .033, V = .12$) than non-RMD students.

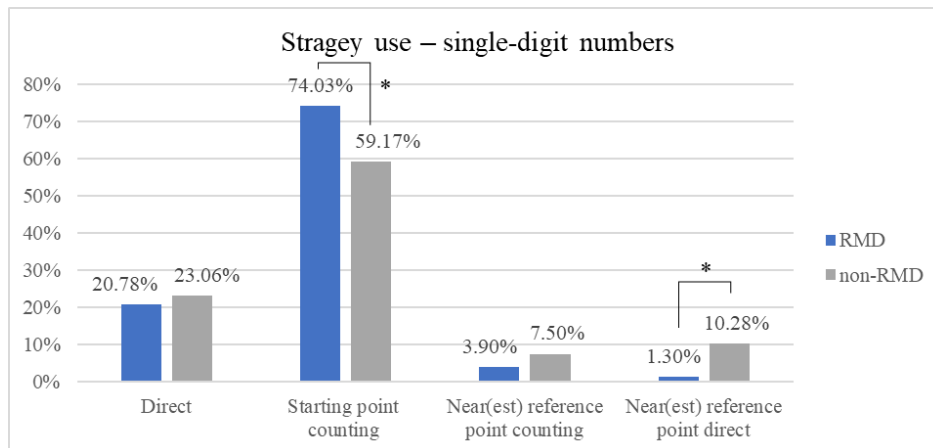


Figure 5: Students’ strategy use for single-digit numbers (* $p < .05$)

Also, for two-digit numbers (Fig. 6), chi-square test showed significant differences in the use of strategies between students with and without RMD: $\chi^2(5) = 24.80, p < .001, V = .25$. Cell tests showed that RMD students used strategy 2 ‘starting point use and counting’ significantly more often ($\chi^2(1) = 23.33, p < .001, V = .24$) and strategy 4 ‘use of near(est) reference point and counting’ significantly less often ($\chi^2(1) = 4.32, p = .038, V = .10$) than non-RMD students.

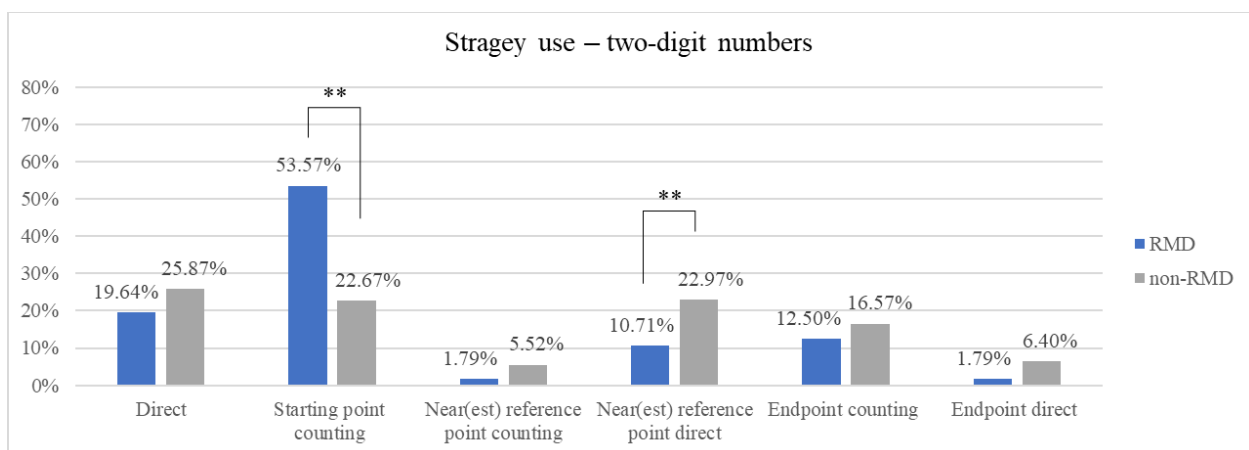


Figure 6: Students’ strategy use for two-digit numbers (** $p < .001$)

DISCUSSION AND CONCLUSION

The aim of this study was to investigate if RMD first graders differ from non-RMD first graders in locating numbers on a marked NL from 0–20. To pursue this aim, we analyzed students’ success rates, response times, and strategy use based on ET data.

Our analyses indicate that success rates differed significantly (medium to large effects) between first graders with and without RMD for locating single-digit numbers and two-digit numbers on a marked NL from 0–20. RMD students made significantly more

errors. This is in line with findings showing that low accuracy in empty NL tasks in six-year-old children can hint at later MD (e.g., Bull et al., 2021). The analyses of response times showed no significant differences, i.e., students with and without RMD needed a similar amount of time to solve the tasks. However, there were significant differences in the use of strategies (small to medium effects). RMD first graders used significantly more counting strategies from the starting point of the NL and used less often the near(est) reference points than non-RMD first graders. This is in line with previous studies on empty NL tasks showing that children's use of reference points (other than the starting point) relates to their experience with numbers and that, depending on children's familiarity with the number range, reference points might need to be labelled for students to be able to use them (Peeters et al., 2017). This could explain the lack of significant differences in response times, despite the more efficient use of strategies of non-RMD first graders. For children of this age, more efficient strategies (where, e.g., reference points are used instead of counting strategies) are apparently not (yet) faster. The use of efficient strategies at this age seems to be demanding, possibly because they require many cognitive resources. Yet, these strategies lead to correct answers more often. Figure 7 shows an example for a 'starting point use and counting' strategy, and an example for 'use of near(est) reference point', whose time on the task was longer than for the counting example on the left-hand side.

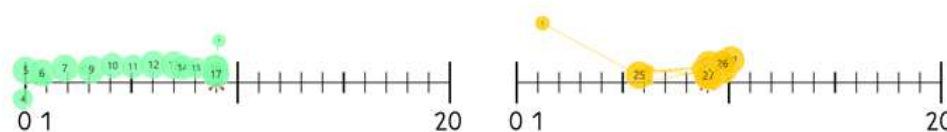


Figure 7: Left: Starting point use and counting (6.99 sec);
Right: Use of near(est) reference point (9.06 sec)

In addition, our results show that counting strategies (from the starting point of the NL) are not yet as conspicuous in RMD children in first grade, as they do not take significantly longer than those of non-RMD children, who tend to use more efficient strategies.

A possible limitation of our study is the relatively small number of RMD students. The RMD students showed great difficulty in solving the NL tasks, so that relatively fewer tasks (especially with two-digit numbers) could be included in the analyses of response times and strategies than for non-RMD students. Further research should investigate if our results can be generalized to a larger number of RMD students. Nevertheless, this study provides insights into the solving of tasks on the marked NL – an important tool in mathematics education that can contribute to the development of number sense – by first graders with and without RMD.

Acknowledgement

This project has received funding by the Erasmus+ grant program of the European Union under grant agreement No 2020-1-DE03-KA201-077597.



Erasmus+
Enriching lives, opening minds.

References

- Bull, R., Lee, K., & Muñez, D. (2021). Numerical magnitude understanding in kindergartners: A specific and sensitive predictor of later mathematical difficulties? *Journal of Educational Psychology*, 113, 911–928.
- Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, 51, 1484–1525.
- Friso-van den Bos, I., Kroesbergen, E. H., Van Luit, J. E. H., Xenidou-Dervou, I., Jonkman, L. M., Van der Schoot, M., & Van Lieshout, E. C. D. M. (2015). Longitudinal development of number line estimation and mathematics performance in primary school children. *Journal of Experimental Child Psychology*, 134, 12–29.
- Geary, D. C. (2011). Consequences, Characteristics, and Causes of Mathematical Learning Disabilities and Persistent Low Achievement in Mathematics. *Journal of Developmental & Behavioral Pediatrics* 32, 250–263.
- Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford UP.
- Moser Opitz, E., Freseemann, O., Prediger, S., Grob, U., Matull, I., & Hußmann, S. (2017). Remediation for students with mathematics difficulties: An intervention study in middle schools. *Journal of Learning Disabilities*, 50, 724–736.
- Peeters, D., Sekeris, E., Verschaffel, L., & Luwel, K. (2017) Evaluating the effect of labeled benchmarks on children's number line estimation performance and strategy use. *Frontiers in Psychology*, 8, 1082.
- Simon, A. L., Asghari, P., Lilienthal, A. J., & Schindler, M. (2023). Strategy use in number line tasks of students with and without mathematical difficulties: A study using eye tracking and AI. In M. Ayalon, B. Koichu, R. Leikin, L. Rubel., & M. Tabach (Eds.), *Proceedings of the 46th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 211–218). PME.
- Simon, A. L., & Schindler, M. (2022). Strategy use in number line tasks of students with mathematical difficulties: An eye-tracking study. In C. Fernández, S. Llinares, A. Gutiérrez, & N. Planas (Eds.), *Proceedings of the 45th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 3–10). PME.
- Teppo, A., & van den Heuvel-Panhuizen, M. (2014). Visual representations as objects of analysis: The number line as an example. *ZDM Mathematics Education*, 46, 45–58.
- van't Noordende, J. E., van Hoogmoed, A. H., Schot, W. D., & Kroesbergen, E. H. (2016). Number line estimation strategies in children with mathematical learning difficulties measured by eye tracking. *Psychological Research*, 80, 368–378.
- Von Aster, M. G., Bzufka, M. W., & Horn, R. R. (2009). *ZAREKI-K Neuropsychologische Testbatterie für Zahlenverarbeitung und Rechnen bei Kindern – Kindergartenversion*. Pearson Assessment & Information.
- Walter, J. (2020). Ein Screening-Verfahren zur Prognose von Rechenschwierigkeiten in der Grundschule. *Zeitschrift für Heilpädagogik*, 71, 238–253.

TEACHER LEADERS' NOTICING IN MODELLING AND ARGUMENTATION IN THE MATHEMATICS CLASSROOM

Horacio Solar^a, Florencia Gómez^a, María Victoria Martínez^b, Andrés Ortiz^c,
Arriagada. Victoria^a

^aPontificia Universidad Católica de Chile, ^bUniversidad de O'Higgins. ^cUniversidad Católica de la Santísima Concepción

This study explores how teacher leaders recognize and interpret critical instructional events within a professional development program centered on argumentation and modelling. Building on existing literature that highlights the significance of teacher support, we examine how teacher leaders' "noticing" practices extend beyond student thinking to encompass nuanced interpretations of classroom interactions. Specifically, this research investigates how teacher leaders identify and analyze critical events related to argumentation and modelling in mathematics education. The findings are presented in two areas: first, the identification of critical events involving the management of argumentation and modelling, and second, the interpretation of these practices.

INTRODUCTION

An increasing number of countries have undertaken significant reforms to their mathematics curricula to help students develop mathematical competencies. These reforms are grounded in the mathematical competencies' framework (Niss & Højgaard, 2019) that has gained international prominence, particularly through the influence of PISA reports. Of all the mathematical competencies promoting these changes, several countries in South and Central America have incorporated argumentation or modelling into their school mathematics curricula: Chile (MINEDUC, 2016), Brasil (Ministério da Educação e Cultura. Secretaria da Educação Básica. 2017), and Cuba (Ministerio de Educación, 2020).

Literature on argumentation in the mathematics classroom comprises several approaches. One such approach, which is commonly accepted in educational research, indicates that the purpose of argumentation is to convince oneself and others of the validity of a line of reasoning (Ayalon & Hershkowitz, 2018; Van Eemeren et al., 2013). Argumentation is an activity performed to convince an audience through claims, evidence, backing, and rebuttals (Cervantes-Barraza et al., 2017; Toulmin, 2003). In the mathematics classroom, argumentation promotes a culture where knowledge construction is regarded as a situated, critical, and reflective activity involving group participation: when all of this occurs, authors speak of collective argumentation (Conner et al., 2014; Krummheuer, 1995).

In mathematical modelling, it is essential to establish an interaction between real-world problems and the mathematical domain, recognizing that real-world problems are inherently complex, while mathematical models serve as simplified representations of this complexity (Blum & Niss, 1991). Modelling involves a sequence of iterative cycles, including simplification, mathematization, working with mathematics, interpretation of results, and validation of the model (Blum & Borromeo-Ferri, 2009; Maaß, 2006). Empirical studies have consistently shown that students encounter significant challenges when engaging in mathematical modelling tasks. Consequently, teacher interventions that support students' transitions across the different phases of the modelling process are critical to fostering their understanding and success (Blum & Borromeo-Ferri, 2009).

The notion of argumentation and modelling in the curriculum can prompt the emergence of divergent professional views among teachers, who may also disagree on how to promote them in the mathematics classroom. If such curricular reforms are to be successful, teachers must develop professional competencies for fostering argumentation and modelling in the school (Fernández et al., 2018; Mason, 2002). To such an end, noticing has emerged as one of the key professional competencies for teachers.

To date, there are several conceptualizations of professional noticing (Blömeke et al., 2015; Jacobs et al., 2010, Santagata & Yeh, 2016). Focusing on the original van Es and Sherin's understanding, noticing has been conceptualized as the ability to identify noteworthy events of classroom interactions and interpret them using one's knowledge and experiences to make sense of what is observed, making connections between what is noticed and broader principles of teaching and learning (van Es & Sherin, 2021).

Literature on teacher noticing in mathematics teaching situations (e.g. Amador, 2017; Fernández et al., 2020; Jacobs et al., 2010) shows that many studies have focused on noteworthy or critical events related to student mathematical thinking. Nevertheless, in contexts where educational purposes are broader, such as the development of mathematical competencies like argumentation and modelling, it can be especially relevant to include critical events that lie beyond the student sphere, such as teacher actions or teacher-student interactions.

Research shows successful teachers' professional development is usually *collaboratively* promoted (Schoenfeld, 2011). Collaboration with peers and participation in learning communities play a critical role in fostering reflective practices grounded in one's teaching experiences. Internationally, professional development initiatives focusing on mathematics teacher noticing increasingly incorporate Teacher Leaders (TLs). This trend is supported by research demonstrating the positive impact of TLs on teacher practices (Witherspoon et al., 2021). Professional development programs that include TLs and emphasize teacher noticing can empower teachers to effectively promote argumentation and modelling in mathematics classrooms (Solar et al., 2023). Based on this context, our research question is: How

do Teacher Leaders notice teacher support for argumentation and modelling in the mathematics classroom?

METHODOLOGY

This study contributes to a broader project aimed at characterizing noticing among teachers supported by TLs in learning to foster argumentation and modelling within classrooms. To accomplish the goals of the overarching initiative, we are implementing a program involving TLs in Chile. The program for TLs comprises three components: argumentation and modelling, noticing, and coaching skills. In this study, we characterize teacher leaders' noticing of critical events in teacher support for argumentation and modelling in the mathematics classroom, analyzing a portion of the Noticing component. This study's Teacher Professional Development program (TPD) involved ten TLs from two large urban regions in Chile. The TPD program included a phase where TLs coached teachers of students aged 6 to 11 from elementary and 12 to 14 from lower secondary levels. The program spanned from the second semester of 2023 to the first semester of 2024, during which TLs practiced their coaching skills. Following this, TLs engage in a 3-semester coaching process with peer teachers, running from the second semester of 2024 to the second semester of 2025, to develop teacher noticing to foster argumentation and modelling in mathematics classrooms.

The teachers participating in the TPD program were selected based on their prior excellence in educational experiences centered on argumentation or modelling, or their completion of a graduate program in mathematical education, making them qualified to support other teachers in this project.

The participants in this study are four of the ten TLs who participated in the TDP program. The other six teachers did not participate on the activity in which the data was collected and therefore were not considered in this study.

The data for this study come from the noticing TPD component that had a focus on argumentation and modelling and comprised four sessions. In the first two sessions TLs analyzed the students' thinking in the context of a learning situation on argumentation and modelling observed in a video. They answered questions such as: what draws your attention in this video? How are students using the mathematics? What challenges are they facing? and once they have discussed their answers, they selected critical events and interpreted them. For this interpretation the TLs had analytical tools such as van Es' scale (2011) that describes 4 levels of noticing as resources. The instrument served as a guiding tool for analyzing peer support. The third session aimed to have the TLs analyze critical events in a learning situation involving argumentation, as presented in a video. Using an argumentative orchestration framework (Author, 2021), the TLs collaboratively identified and interpreted these critical events. The fourth session, which provided the data for this study, focused on analyzing critical events in a learning situation involving modelling, also presented in a video. During this session, the group work of the four participating TLs was video-recorded and transcribed for analysis.

The TLs were given a learning scenario designed for 3rd-grade students (7–8 years old), featuring a mathematical task. In this task, students were asked to build a model of a walkway using interlocking cubes, supported by two toy vehicles provided by the teacher: a car and a bus. This learning situation was considered appropriate for enabling the TLs to recognize critical events related to argumentation and modelling.

The discussion among the four TLs during the fourth session was transcribed. All episodes in which the TLs identified and interpreted critical events related to the argumentative orchestration of the teacher in the video or to teaching strategies for promoting modelling were coded.

An open coding process was employed to analyze the teachers' responses (Creswell, 2011), with a focus on identifying the main themes from their discussions in response to questions 1 and 2—specifically, the identification of critical events and the interpretation of those events. For argumentation, particular attention was given to teaching actions, while for modelling, emphasis was placed on the stages of the cycle.

FINDINGS

In the analyzed session, the TLs answered each of the questions in the activity. For question 1, what draws your attention to how the teacher manages the students' math work? Why do you think the events that caught your attention occurred? Question 2: What draws your attention to how the teacher manages the sharing of results and procedures? Why do you think the events that caught your attention occurred? They completed a table identifying the critical events and the interpretation (see table 1).

Table 1

Critical events about teacher management of the learning situation identified by TLs

Critical events	Identification	Interpretation
Managing student work	a) The use of counter-questions and the use of questions from other groups, without validating responses and interventions, supported by specific material	The use of counter-questions is not to give the students the answer but rather to allow them to discuss and validate their opinions. In addition, it asks them to check their ideas using the concrete material. With this, they can also refute the answers of their classmates.
	b) Involve students from other groups by including them in a particular group's discussion	Focus on ideas and discussion arising from the students themselves, with the teacher taking a mediating role to create instances of argumentation

Managing student sharing of their work	c) That the sharing is different from the traditional ones since it calls a representative of each group to discuss the procedures or ideas of a particular group avoiding validating those procedures or ideas	<p>So that each representative can take the discussion to their groups.</p> <p>This is so that members of each group can observe what their classmates did and compare it with their own models, analyzing whether it answers the problem.</p> <p>The teacher does not validate the results, but she asks guiding questions to guide the procedures.</p>
--	---	--

The analysis of the teacher leaders' productions and their interventions revealed two findings related to our research question. The first finding refers to identifying critical events associated with the management of argumentation and modelling, and the second finding is the interpretation of these critical events.

Below is an extract that shows how the TLs identify and interpret the making of counter-questions and failure to validate as part of the argumentative orchestration (critical event a).

04:05 FINA: In the first one, when we wrote down the use of counter questions, we could also add she never, or I don't know if, but I didn't notice her, she doesn't validate any answer as good or bad, but instead, she asks counter questions, it includes the questions of children from other groups; above all, imagine yourself the girl who insisted a lot on the question, because the girl asks: "what happens if there are two buses? They won't fit", so it's like there's all that discussion between the same students.

04:55 ÁNGELES: Ok, here it's without validating. So, she doesn't validate.

05:06 FINA: Yes, she doesn't, I mean, the answers from the students.

DISCUSSION

Two key findings are presented to address the objective of this study to characterize the noticing of Teacher Leaders (TLs) in teaching support to promote argumentation and modelling in the mathematics classroom: the first is in relation to the identification of critical events associated with the management of argumentation and modelling, and the second with a focus on the interpretation of argumentation and modelling. Regarding the first finding, it is striking that in a training session in the MADC program whose purpose was for TLs to analyze critical events in a learning situation about

modelling, TLs focused mainly on the management of argumentation over the management of modelling.

We believe that these findings are particularly novel, as much of the research on noticing follows the approach of Jacobs et al. (2010) and van Es & Sherin (2002), focusing on critical events related to students' mathematical thinking (Amador, 2017; Fernández et al., 2020; Jacobs et al., 2010). While recent studies have addressed noticing in argumentation (Ayalon, & Nama, 2024), and noticing in modelling (Alwast & Vorhölter, 2022; Cai et al., 2022), these studies also focus on noticing in students' mathematical thinking. Our study's findings extend this research by characterizing TLs' noticing of critical events focused on teachers' ability to manage argumentation and modelling in the mathematics classroom. While Rotem & Ayalon's (2022) study identified several types of analysis that a teacher can conduct, such as interpreting student thinking, interpreting teacher responses, and responding to student thinking, our study goes further by identifying critical events associated with specific teacher actions for managing argumentation and modelling. This characterization of critical events is especially important when pursuing broader educational goals, such as developing argumentation and modelling.

Regarding the second finding on interpreting critical events, another aspect that draws our attention is the interrelations that the TLs make between argumentation and modelling. During the TDP, the TLs analyzed situations of argumentation and modelling. In the analyzed session of Noticing with a focus on argumentation and modelling, we observed that the TLs make interrelated connections to interpret the identified critical events, not differentiating between teaching actions more typical of argumentation and modelling. There are relatively recent studies (Tekin-Dede, 2019) accounting for how arguments are constructed in the modelling cycle in students.

The findings of our study lead us to believe that the leading teachers who participate in the TDP program are developing the necessary tools to support other teachers in improving their capacity to notice in argumentation and modelling. However, they must continue developing the capacity to differentiate the teaching strategies specific to argumentation and modelling.

References

- Amador, J. M. (2017). Preservice teachers' video simulations and subsequent noticing: a practice-based method to prepare mathematics teachers. *Research in Mathematics Education*, 19(3), 217–235. <https://doi.org/10.1080/14794802.2017.1315317>
- Alwast, A., & Vorhölter, K. (2022). Measuring pre-service teachers' noticing competencies within a mathematical modelling context – an analysis of an instrument. *Educational Studies in Mathematics*, 109(2), 263–285. <https://doi.org/10.1007/s10649-021-10102-8>
- Ayalon, M., & Hershkowitz, R. (2018). Mathematics teachers' attention to potential classroom situations of argumentation. *The Journal of Mathematical Behavior*, 49, 163–173. <https://doi.org/10.1016/j.jmathb.2017.11.010>

- Ayalon, M. & Nama, S. (2024) Secondary school mathematicsteacher-perceived factors involved in argumentation: an emerging framework, *Research in Mathematics Education*, 26:1, 193-214, DOI: 10.1080/14794802.2022.2156585
- Blum, W., & Niss, M. (1991). Applied mathematical problem solving, modelling, applications, and links to other subjects: State, trends and issues in mathematics instruction. *Educational Studies in Mathematics*, 22(1), 37–68. <https://doi.org/10.1007/bf00302716>
- Blum, W. & Borromeo-Ferri, R. (2009). Mathematical modelling: can it be taught and learnt? *Journal of Mathematical Modelling and Application*, 1(1), 45-58.
- Blömeke, S., Hoth, J., Döhrmann, M., Busse, A., Kaiser, G., & König, J. (2015). Teacher Change During Induction: Development of Beginning Primary Teachers' Knowledge, Beliefs and Performance. *International Journal of Science and Mathematics Education*, 13(2), 287–308. <https://doi.org/10.1007/s10763-015-9619-4>
- Cai, J., LaRochelle, R., Hwang, S., & Kaiser, G. (2022). Expert and preservice secondary teachers' competencies for noticing student thinking about modelling. *Educational Studies in Mathematics*, 109(2), 431-453.
- Cervantes-Barraza, J., Cabañas-Sánchez, G. & Ordoñez-Cuastuma. (2017). El poder persuasivo de la refutación en argumentaciones colectivas. *Bolema*, 31 (59), 861–879.
- Conner, A. M., Singletary, L. M., Smith, R. C., Wagner, P. A. & Francisco, R. T. (2014). Teacher support for collective argumentation: A framework for examining how teachers support students' engagement in mathematical activities. *Educational Studies in Mathematics*, 86(3), 401–429.
- Creswell, J. (2011). *Educational research: planning, conducting, and evaluating quantitative and qualitative research*. Pearson.
- Fernández, C., Sánchez-Matamoros, G., Valls, J. & Callejo, M.L. (2018). Noticing students' mathematical thinking: characterization, development and contexts. *Avances de Investigación en Educación Matemática (AIEM)*, 13, 39–61. <https://doi.org/10.35763/aiem.v0i13.229>
- Jacobs, V. R., Lamb, L. L. C., & Philipp, R. A. (2010). Professional noticing of children's mathematical thinking. *Journal for Research in Mathematics Education*, 41(2), 169–202. <https://doi.org/10.5951/jresmetheduc.41.2.0169>
- Krummheuer, G. (1995). The ethnography of argumentation. En P. Cobb y H. Bauersfeld (Eds.), *The emergence of mathematical meaning: Interaction in classroom cultures* (pp. 229–269). Hillsdale, NJ: Lawrence Erlbaum.
- Maaß, K. (2006). What are modelling competencies? *ZDM – Mathematics Education*, 38(2), 113-142. <https://doi.org/10.1007/BF02655885>
- Mason, J. (2002). *Researching your own practice. The discipline of noticing*. Routledge Falmer
- Ministério da Educação e Cultura. Secretaria da Educação Básica. (2017). *Base Nacional Comum Curricular*. <http://basenacionalcomum.mec.gov.br/#/site/inicio>

- Ministerio de Educación. (2020). *Adaptaciones curriculares para el curso escolar 2020-2021: Educación Primaria y Especial*. Editorial Pueblo y Educación.
- MINEDUC (2016). *Bases curriculares 7º básico a 2º medio*. Santiago: Autor.
- Niss, M. & Højgaard (2019). Mathematical competencies revisited. *Educational Studies in Mathematics*, 102, 9–28
- Rotem, S. H., & Ayalon, M. (2022). Building a model for characterizing critical events: Noticing classroom situations using multiple dimensions. *The Journal of Mathematical Behavior*, 66, 100947.
- Santagata, R., & Yeh, C. (2016). The role of perception, interpretation, and decision making in the development of beginning teachers' competence. *ZDM – Mathematics Education*, 48(1–2), 153–165. <https://doi.org/10.1007/s11858-015-0737-9>
- Schoenfeld, A. H. (2011). Toward professional development for teachers grounded in a theory of decision making. *ZDM*, 43, 457–469. <https://doi.org/10.1007/s11858-011-0307-8>.
- Solar, H., Ortiz, A., Aravena, M., & Goizueta, M. (2023). Relaciones entre la argumentación y la modelación en el aula de matemáticas. *Bolema: Boletim de Educação Matemática*, 37(76), 500–531. <https://doi.org/10.1590/1980-4415v37n76a07>
- Tekin-Dede, A. (2019). Arguments constructed within the mathematical modelling cycle. *International Journal of Mathematical Education in Science and Technology*, 50(2), 292–314. <https://doi.org/10.1080/0020739X.2018.1501825>
- Toulmin, S. (2003). *The uses of argument*. Cambridge University Press (Second ed.).
- Van Eemeren, F. H., Grootendorst, R., Johnson, R. H., Plantin, C., & Willard, C. A. (2013). *Fundamentals of argumentation theory: A handbook of historical backgrounds and contemporary developments*. Routledge
- van Es, E. A. (2011). A framework for learning to notice student thinking. In Sherin, M. G., Jacobs, V., Philipp, R. (Eds.), *Mathematics teacher noticing: Seeing through teachers' eyes* (pp. 134–151). Routledge.
- van Es, E. A., & Sherin, M. G. (2002). Learning to Notice: Scaffolding New Teachers' Interpretations of Classroom Interactions. *Journal of Technology and Teacher Education*, 10, 571–596.
- Van Es, E. A., & Sherin, M. G. (2021). Extending on prior conceptualizations of teacher noticing. *ZDM Mathematics Education*, 53(1), 17–27. <https://doi.org/10.1007/s11858-020-01211->
- Witherspoon, E. B., Ferrer, N. B., Correnti, R. R., Stein, M. K., & Schunn, C. D. (2021). Coaching that supports teachers' learning to enact ambitious instruction. *Instructional Science*, 49(6), 877–898.

PERCEIVED DIFFICULTY IN MATHEMATICS EDUCATION

Camilla Spagnolo

University of Ferrara

Mathematics education researchers have widely studied the possible characteristics that influence the difficulty of solving a task. This study focuses on the perspective of students and teachers and their perceived difficulty, which have not been equally addressed. Difficulty in an absolute sense and perceived difficulty are different but intertwined: task characteristics can shape students' and teachers' perceptions, and these perceptions of difficulty can affect the solver's behaviour, influencing the accuracy of the response. Our study focuses on the factors influencing secondary school students' and teachers' evaluations of task difficulty. We describe the process that led to the development of macro-categories for perceived difficulty, highlighting both the study's limitations and potential directions for future research.

INTRODUCTION

The challenge of difficulty in mathematics has been extensively examined within the field of Mathematics Education. Difficulties can be classified into different categories depending on the focus: difficulties related to mathematics as a discipline or difficulties related to students in terms of deficits or their specific relationship with mathematics (e.g. Zan et al., 2006).

However, the issue of *perceived difficulty* (PD) and its underlying causes have not received equal attention. In particular, PD and its causes have not been addressed in the same way. Within the field of mathematics education research, there is no shared definition of PD, and few studies have explored the factors related to this issue (Eccles & Wigfield, 2020; Doz et al., 2023). It is therefore necessary to first clarify what is the meaning of PD. In this research, we consider PD in relation to the process of solving a mathematical task. During the process of solving a task, a student may encounter various difficulties, which can stem from both the student's own characteristics (such as abilities, knowledge, beliefs, and attitudes) and the characteristics of the task itself (such as the wording or the mathematical content involved). These features of the mathematical task are also the ones that can influence the student's perception of the task and, consequently, their PD. Therefore, while closely related, difficulty and PD are two distinct aspects (Spagnolo & Saccoletto, 2023). We argue that one of the main differences is that PD can influence students' behaviour when approaching a mathematical task. In addition, this research examines the difficulties associated with mathematical tasks, with particular emphasis on the PD experienced by students and teachers *after* facing a task. The literature differentiates between difficulty before and after a task has been attempted or solved, a distinction that is crucial for understanding

also PD. However, this study specifically focuses on PD after students and teachers have faced the task.

The research began in 2021 during the pandemic period, with a study presented at ICME14 published in an extended paper (Saccoletto & Spagnolo, 2022). Here, the aim is to highlight the process that led to the identification of the factors, the development of categories for analyzing PD, and the description of these categories. Early phases of this study have been presented at international conferences such as ICME14, CERME13, MAVI29, ICME15, PME47, MAVI30, and CERME14. Each of these studies provided key elements to the identification of these macro-categories.

THEORETICAL BACKGROUND

In Mathematics Education, there is no shared definition of PD of a task, although it is generally recognized that PD differs from a task's inherent difficulty, which is often determined retrospectively based on students' results (Mehrens & Lehmann, 1991). Instead, research in cognitive psychology has analysed the theme of subjective difficulty and its perception since the 1990s, resulting in various definitions and expressions to characterise the concept (Eccles & Wigfield, 2020; Doz et al., 2023).

First of all, generally speaking, the perception of the difficulty of a task is sometimes considered dependent on the cognitive load that the task implies (Hellmann & Nuckles, 2013). Sometimes, PD has been regarded as a different manifestation of self-efficacy rather than a complementary aspect to it, but this unification does not seem justifiable (Eccles & Wigfield, 2020). One of the concepts closer to the PD as we mean it is the "feeling of difficulty" (FOD), progressively defined by Efklides and colleagues in different studies between the 1990s and 2011. Initially, the FOD was simply described as one of "the closest estimates of feelings of difficulty" (Efklides et al., 1998). The authors acknowledged the impossibility of directly measuring feelings but considered it possible instead to have people rate the difficulty of a task. In a second moment, Efklides and colleagues analysed the various causes of FOD, defining it as a "metacognitive experience that monitors cognitive processing as it takes place" (Efklides & Touroutoglou, 2010, p. 272). FOD is related to PD, yet they are conceptually distinct, as FOD has an "experiential nature" while PD arises from metacognitive judgments based on a conscious recall of self-knowledge and task-related information. Obviously, individuals may draw upon their metacognitive knowledge about the task and themselves to interpret the feelings they experience (Efklides & Touroutoglou, 2010). Taking into account the differences, in this paper we refer to PD, considering the characteristics that make it similar to FOD and adapting them to our context, that is Mathematical Education research. It is worth mentioning the synthesis presented by Doz et al. (2023), stating that the feeling of task difficulty is metacognitive in nature because it derives from monitoring the activity of a developing task processing and the consciousness of this process influences self-regulation, effort, affect, and strategy use.

In order to define the aforementioned macro-categories that describes PD, we draw upon theories and constructs used in Mathematical Education research. Firstly, we considered as a reference to analyse students' answers the *multidimensional model of attitude* (Di Martino & Zan, 2010), which takes into account the perceived competence, the emotional aspects and the vision of mathematics. In addition to that, the aspect of metacognition was also considered, as it is connected to students' decision-making process when addressing a problem, a process influenced by personal beliefs and values (Radmehr & Drake, 2017). We also considered the concept of *expert blind spot* (Nathan & Petrosino, 2003), an expression used to identify teachers' lack of awareness of students' understanding and difficulties associated with great content knowledge. Moreover, the experts may even be unaware of this blind spot deriving from expertise itself. Referring to the concepts of pedagogical content knowledge and subject matter knowledge (Ball et al., 2008), the authors argue that teachers showing an expert blind spot may have both of them but, when applying this knowledge to "a specific area of mathematics, such as algebra instruction, those bodies of knowledge come into conflict" (Nathan & Petrosino, 2003). The concept of expert blind spot can be another hypothesis to explain the different PD that students and teachers experience.

METHODOLOGY

This section will describe the different phases that allowed for the development and definition of the macro-categories. These macro-categories allow to describe factors influencing students' and teachers' PD. Specifically, the first phase, presented at ICME14 (Saccoletto & Spagnolo, 2022), aimed to investigate the correlation between PD and students' ability to answer the task correctly. Then, the second phase, presented at CERME13 (Spagnolo & Saccoletto, 2023), ICME15 and PME47 (Nicchiotti & Spagnolo, 2024), made it possible to clarify the categories that were used to classify the factors influencing students' PD. These categories were determined using Constructive Grounded Theory (CGT), employing an inductive method and directly drawing on the data collected (Glaser & Strauss, 1967). Phases 3 (presented at MAVI29, MAVI30, and CERME14) comparing the factors that influence PD of teachers. All the research phases involved secondary school students and teachers.

First phase: correlation between PD and students' ability to answer the task

In this first qualitative phase, 79 Italian students from a high school from humanities curriculum were involved: two grade 9 classes and two grade 10 classes. The students completed an online questionnaire, followed by in-depth interviews. The questionnaire was administered via Google Forms (using the classroom computers), and a researcher was present during the administration. The questionnaire included two argumentative algebraic tasks related to literal calculus (Figure 1). Specifically, Task 1 is a multiple-choice question that requires the recognition of a correct argument, while Task 2 is an open-ended question that requires the production of an argument; both tasks were selected from previous national INVALSI tests, as they are statistically validated (Lazarsfeld, 1958).

Task 1

Antonio states that « $4n-1$ is always a multiple of 3».
Is Antonio right?
In the following table select the only argumentation that justifies the right answer

Antonio is right...	Antonio is not right...
A. <input type="checkbox"/> because $4n-1=3n$	C. <input type="checkbox"/> because $4n-1$ is always odd
B. <input type="checkbox"/> because if $n=4$ then $4n-1=15$	D. <input type="checkbox"/> because if $n=3$ then $4n-1=11$

Task 2

Marco states that, for every natural number n greater than 0, $n^2 + n + 1$ is a prime number.
Is Marco right?

Choose one of the two answers and complete the sentence.

☐ Marco is right because

.....

☐ Marco is not right because

.....

Figure 1: Task 1 and Task 2 (original texts from www.gestinv.it., translation provided by the authors).

After solving each task, students evaluate its difficulty (explaining the reasons for their evaluation) and answer other questions related to the PD of the specific task. Finally, the last section of the questionnaire was more general, with questions aimed at clarifying the underlying factors that might influence the PD. The interviews were conducted remotely via Google Classroom (Meet applet) with the aim of helping to categorize some of the responses provided by the students. Among the works addressing the need to develop theoretical frameworks on affect, we specifically refer to the study by Di Martino and Zan (2010) on attitude, as we recognized some analogies with their study in reading and analyzing the responses of our students.

Second phase: development and generalization of the categories

The second phase of the study (Spagnolo & Saccoletto, 2023) involved 148 students from schools of different types (five grade 9 classes, and two grade 10 classes). The students completed the same online questionnaire (described in the previous section), followed by in-depth, face-to-face interviews aimed at facilitating the categorization of the responses provided by the students. These responses were categorized using CGT (Charmaz, 1994), with particular attention to the open-ended questions in the questionnaire. The categories of analysis were developed from the specific cases, but conclusions were drawn from all the database (which consisted of descriptive data provided by the students). The categories that emerged from the analysis were compared with the categories of the theoretical framework proposed by Di Martino and Zan (2010), highlighting differences with PD. To analyze and emphasize these differences, focus groups were conducted with the same students. Significant regularities in the data collection proved valuable insight in constructing the categories of analysis.

Third phase: comparison factors influencing PD of students and teachers

Finally, the third phase of the study (Nicchiotti & Spagnolo, 2024) involved teachers, specifically 49 secondary mathematics teachers. The individual characteristics of the teachers varied considerably, particularly in terms of the type of school and curriculum in which they teach, years of experience, and educational background. As in the previous phases, the teachers completed the online questionnaire (with respect to the

same tasks evaluated by the students), followed by in-depth interviews conducted via Google Meet. The qualitative analysis of the data was grounded in the factors characterizing the students' PD, which were compared with factors expressed by the teachers.

RESULTS AND DISCUSSION

We used Excel to calculate mean values and represent the data graphically, while NVIVO 12 was used to conduct the qualitative analysis of the responses expressed by students and teachers.

First phase: correlation between PD and students' ability to answer the task

Analysis highlights that student's PD does not appear to be related to the ability to answer a question correctly but seems to encompass broader metacognitive and self-perception factors. From the qualitative analysis of the questionnaire, it also emerges that there is no correspondence between the PD level selected during the evaluation of individual tasks (in response to "On a scale from 1 to 10, how difficult did you find Task 1/Task 2?") and the level selected during the comparison of the two tasks (in response to "Compare the two tasks. Which one did you find more difficult?"). In particular, students express different level of PD in relation to a single task or in the comparison of multiple tasks (see Figure 1 for Task 1 and 2). Student interviews suggest that this choice is influenced by factors related to attitudes and beliefs. In our view, this may be evidence of students' difficulty in evaluating a task or may indicate that students consider different factors when assessing individual tasks versus comparing them.

Second phase: development and generalization of the categories

The aim of this phase was to further investigate the motivations that lead students to choose one level of difficulty over another. An initial analysis highlighted several categories related to the main aspects mentioned by students in their responses. We considered these categories, re-examined the responses, and worked to unify, compare, and clarify the main aspects. The analysis provide five main macro-categories (Spagnolo & Saccoletto, 2023): *Resolution Strategy*, *Capabilities and Experience*, *Emotions*, *Task Formulation*, and *Self Consideration*.

Resolution Strategy category, grouped responses where students explicitly referred to the type of strategy or process they believed was necessary to solve the task. The second category, *Capabilities and Experience*, is the most prevalent and pertains to responses that refer to the abilities or competencies students perceive and to previous experiences that influence their PD toward the task. This category also includes students who explicitly state that they are unfamiliar with this type of task, reinforcing the idea that a problem is easier if it resembles something already known. Additionally, this category includes responses referring to what students can (cannot) do or (do not) know. The focus in these responses is on students' self-perception (either in general or regarding the tasks). The *Emotions* category refers to students explicitly considering

their emotions when justifying the level of PD chosen. The fourth category, *Task formulation*, represents considerations about the task's structure, especially in relation to the wording. The fifth category, *Self Considerations*, refers to students' personal reflections on their success in mathematics.

The categories emerging from the analysis help clarify the main aspects involved when a student expresses her or his PD in relation to mathematical tasks. These categories are not exclusive, and some responses may be classified under more than one category.

We believe that these five categories can be linked to the attitude construct in the sense of Di Martino and Zan (2010). Additionally, the responses to the questionnaire can be related to students' view of mathematics: some students explicitly describe the methods they consider necessary to solve tasks. The Self Considerations category contains responses that express what students believe is necessary for success in mathematics. Finally, even in our case, students refer, more or less explicitly, to their perceived competencies in solving tasks, while also expressing ideas about their perceived knowledge and skills. However, when assigning a PD level to a task, students seem to be influenced by factors more directly related to the task (such as elements in the wording), by factors linked to their attitude or emotions, and by metacognitive aspects (such as the inability to assess their own competencies, knowledge, and skills, or personal reflections on how to improve their performance).

The analysis of students' ratings and answers allowed also to state that there seem to be some differences between boys' and girls' PD. Boys tend to evaluate mathematical tasks as easier than girls do, even regardless of their actual performance solving them (Nicchiotti & Spagnolo, 2024).

Third phase: comparison factors influencing PD of students and teachers

This phase provides an initial overview of the factors influencing teachers' PD, comparing them with those influencing students' PD. The macro-categories defined in Spagnolo & Saccoletto (2023) are useful for analyzing teachers' responses. The factors influencing the PD of both teachers and students are similar, but they have different proportion. Generally, teachers appear to be aware of the reasons behind PD, but sometimes they underestimate them and seem to overlook the emotional aspects related to difficulty. This tendency aligns with the notion that experts, due to their deep knowledge, may struggle to recognize the challenges faced by learners, as their expertise can create a disconnect between their understanding and students' perceptions. In this context, teachers seem to focus more on objective factors, such as the form and content of the task, when assessing its PD. In contrast, students also consider subjective elements, such as experience or self-perception. This may reflect a broader cognitive gap that the expert blind spot theory helps to explain.

CONCLUDING REMARKS

This research gives a first insight into the PD and the factors characterizing it, comparing students' and teachers' perspectives. PD has a strong affective connotation,

as it entails beliefs, self-efficacy and emotions. The main factors used to describe PD of students and teachers in relation to a mathematical task are summarized in the five macro-categories: Resolution Strategy, Capabilities and Experiences, Emotions, Task Formulation, and Self-consideration. Figure 2 shows a visual representation of PD. The eye, symbolizing the personal perspective, is surrounded by the five macro-categories influencing PD, which intersect at the center of the eye. This highlights that the macro-categories are not mutually exclusive, but may overlap (meaning that a student or teacher may consider more than one category when explaining why they perceive a task as difficult or easy).

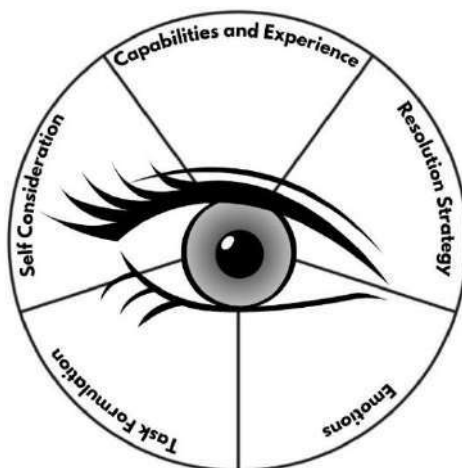


Figure 2: Visual representation of macro-categories that describe PD.

The five macro-categories based on the factors influencing students' PD *after* solving mathematical tasks are essential for providing a definition of PD. However, the validity of these findings could be strengthened considering expanding the study working with a bigger sample and more tasks, of different types and regarding other math topics. Further study will examine the phenomenon from a quantitative perspective.

References

- Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content Knowledge for Teaching: What Makes It Special? *Journal of Teacher Education*, 59(5), 389-407. <https://doi.org/10.1177/0022487108324554>
- Di Martino, P., & Zan, R. (2010). 'Me and maths': Towards a definition of attitude grounded on students' narratives. *Journal of mathematics teacher education*, 13(1), 27-48.
- Doz, E., Cuder, A., Pellizzoni, S., Carretti, B., & Passolunghi, M. C. (2023). Arithmetic Word Problem-Solving and Math Anxiety: The Role of Perceived Difficulty and Gender. *Journal of Cognition and Development*, 1-19. <https://doi.org/10.1080/15248372.2023.2186692>
- Eccles, J. S., & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology*, 61. Art 101859. <https://doi.org/10.1016/j.cedpsych.2020.101859>

- Efklides, A., Papadaki, M., Papantoniou, G., & Kiosseoglou, G. (1998). Individual differences in feelings of difficulty: The case of school mathematics. *European Journal of Psychology of Education*, 13(2), 207-226.
- Efklides, A., & Touroutoglou, A. (2010). Cognitive interruption as an object of metacognitive monitoring: Feeling of difficulty and surprise. In: A. E. & P. Misailidi (Ed.), *Trends and prospects in metacognition research* (pp. 171–208). New York, NY: Springer. https://doi.org/10.1007/978-1-4419-6546-2_6
- Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory. Strategies for qualitative research*. Chicago: Aldine.
- Hellmann, K., & Nuckles, M. (2013). Expert Blind Spot in Pre-Service and In-Service Mathematics Teachers: Task Design moderates Overestimation of Novices' Performance. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 2518–2523).
- Lazarsfeld, P. F. (1958). Evidence and inference in social research. *Daedalus*, 87(4), 99-130.
- Mehrens, W. A., & Lehmann, I. J. (1991). *Measurement and evaluation in education and psychology*. Fort Worth: Holt, Rinehart & Winston.
- Nathan, M. J., & Petrosino, A. (2003). Expert Blind Spot Among Preservice Teachers. *American Educational Research Journal*, 40(4), 905-928. <https://doi.org/10.3102/00028312040004905>
- Nicchiotti, B., & Spagnolo, C. (2024). Gender differences in relation to perceived difficulty of a mathematical task. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.), *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 257–264). PME
- Radmehr, F., & Drake, M. (2017). Exploring students' mathematical performance, metacognitive experiences and skills in relation to fundamental theorem of calculus. *International Journal of Mathematical Education in Science and Technology*, 48(7), 1043-1071.
- Saccolotto, M., & Spagnolo, C. (2022). Students' perceived difficulty of mathematical tasks: an investigation on influencing factors. *Didactica Mathematicae Journal*, 44, 59-79. <https://doi.org/10.14708/dm.v44i1.7181>
- Spagnolo, C., & Saccolotto, S. (2023). How students view the difficulty of mathematical tasks: factors that influence their perceptions. In P. Drijvers, C. Csapodi, H. Palmér, K. Gosztonyi, & E. Kónya (Eds.), *Proceedings of the Thirteenth Congress of the European Society for Research in Mathematics Education (CERME13)* (pp. 1498–1506). Alfréd Rényi Institute of Mathematics and ERME.
- Zan, R., Brown, L., Evans, J., & Hannula, M. (2006). Affect in mathematics education: an introduction. *Educational studies in mathematics*, 63(2), 113–121. <https://doi.org/10.1007/s10649-006-9028-2>

USING KNOWLEDGE ABOUT PROOF PRINCIPLES FOR PROOF VALIDATION AT SECONDARY SCHOOL LEVEL

Femke Sporn, Daniel Sommerhoff, Aiso Heinze

IPN - Leibniz Institute for Science and Mathematics Education, Kiel, Germany

Students are not only expected to construct proofs, but also to be able to check proof validity. So-called proof principles, which can be regarded as a minimal set of criteria for the validity of mathematical proofs, can be used for validation. This study investigated to what extent $N = 456$ students in grades 8 to 11 know about these proof principles and use this knowledge to validate purported proofs. The results show that most students struggled to identify and correctly evaluate disregarded proof principles in invalid purported proofs. Moreover, many students did not show the knowledge that disregarding these principles logically leads to the invalidity of the purported proof. The results suggest that secondary school students do not have adequate knowledge about proof principles, which could explain uncertainties in proof validation.

INTRODUCTION

Proofs are a central part of mathematics and are therefore also integrated into secondary mathematics curricula worldwide. Thus, learners are repeatedly confronted with proofs and related activities. This focus on proofs and related activities is part of mathematics education at school and university (Komatsu et al., 2017). Students are not only supposed to learn how to construct proofs but also how to validate given proofs and, finally, accept or disprove their validity (Stylianides et al., 2017).

Different criteria are used for proof validation, which may vary across communities (Sommerhoff & Ufer, 2019). However, there is a minimal set of criteria – which we call proof principles – that must be fulfilled by argumentations to be accepted as a valid mathematical proof (Heinze & Reiss, 2003; Stylianides, 2007). To validate purported proofs successfully, students need adequate knowledge about these proof principles (Selden & Selden, 2003) and be able to apply it purposefully. Thus, aims of mathematics education is to help learners to (A1) acquire knowledge about the proof principles, (A2) be able to identify proof principles and check their fulfillment in purported proofs, and (A3) know that disregarding (one of) the proof principles, for logical reasons, means that the purported proof cannot be accepted as a valid mathematical proof.

In prior empirical studies, students were often asked to determine whether purported proofs were valid or invalid (e.g., Alcock & Weber, 2005; Selden & Selden, 2003). However, to the best of our knowledge, it has not yet been explicitly investigated whether or to what extent students know about the logical consequences of

disregarding just one of the proof principles in a purported proof and how they use this knowledge for accepting or rejecting the overall proof (A3).

THEORETICAL BACKGROUND

Mathematical Proof and Proof Principles

Proofs are of central importance in mathematics. One challenge in mathematical practice is that the validity of a proof depends on its acceptance within the community. Acceptance criteria differ across various communities, which is why they are also described as socio-mathematical norms for proof (Sommerhoff & Ufer, 2019; Yackel & Cobb, 1996). Thus, standardized and generally accepted criteria for unambiguously determining the validity of mathematical proofs are lacking (Hanna & Jahnke, 1996). However, there is a minimal set of criteria that can be assumed to be necessary in any community for an argument to constitute a valid mathematical proof (e.g., Selden & Selden, 2003). According to Heinze and Reiss (2003), this minimal set of criteria can be classified into three categories of proof principles: *proof scheme*, *proof structure*, and *chain of conclusions*. *Proof scheme* includes criteria for invalid arguments and thus, for example, that a reference to a higher authority is not an adequate argument for a mathematical proof. *Proof structure* means that a proof must prove what it is supposed to prove and includes, for example, that using the assertion as an argument is not sufficient. *Chain of conclusions* means, that each step of a proof can be concluded from the previous steps.

Proof Validation

The validation of proofs aims to verify whether an argumentation can be considered as a proof in a mathematical sense. Given a purported proof, it must be evaluated against the socio-mathematical norms of the community to decide if it can be accepted as a valid proof or not consequentially (Selden & Selden, 2015; Yackel & Cobb, 1996). Although there are no standardized criteria sufficient for verifying the validity of proofs in mathematical practice, proof principles as minimal set of criteria can be used in the sense of necessary conditions in proof validation (Selden & Selden, 2003). Based on the evaluation of these proof principles, the overall proof can be judged as valid or invalid. In particular, if any of the proof principles is disregarded, the purported proof as a whole should not be accepted as a valid mathematical proof.

Proof validation is one of the proof-related activities students have to learn as part of their mathematical education. Empirical studies have repeatedly shown that learners have difficulties in validating proofs (e.g., Alcock & Weber, 2005; Healy & Hoyles, 1998). For example, Selden and Selden (2003) found that many students incorrectly accepted an argument showing the truth of the converse of a target proposition as a valid proof. The result suggests that it is difficult for the students to validate the purported proof correctly regarding the proof principle *proof structure*.

Students' Knowledge About Proof Principles

In order to validate a purported proof, students need to identify the relevant proof principles and then to use them as criteria to judge the validity of the overall proof. Hence, as a prerequisite for proof validation, students need to develop adequate knowledge about the proof principles. Previous empirical studies investigated this knowledge by explicitly asking learners to evaluate possible proof principles without involving any proof-related activity. For example, in Sporn et al. (2021) students had to evaluate the possible proof principle ‘If the most important mathematicians in a field consider a statement to be true, then it can be considered valid, even if there is no complete proof yet.’ (A1). Selden and Selden (2003) used a different approach and investigated knowledge about proof principles through a proof validation task. In their study, students were asked to validate purported proofs that disregarded proof principles. Thus, it was investigated whether the students could identify and correctly evaluate the proof principles in purported proofs (A2 & A3). In this way, Alcock and Weber (2005) also showed that it is difficult for students to correctly identify disregarded proof principles in purported proofs (A2).

Knowledge about proof principles also means that students know about their logical meaning for the validity of mathematical proofs (Selden & Selden, 1995). For example, someone might correctly describe a proof principle but fail to recognize that a purported proof cannot be judged as valid if the proof principle has been disregarded. Such inadequate knowledge about the role of proof principles can also manifest in cases where someone identifies a proof principle as ‘disregarded’ in a purported proof but does not conclude that the proof is invalid.

RESEARCH QUESTIONS

So far, and to the best of our knowledge, it is not known if learners know about proof principles regarding their logical meaning for the validity of mathematical proofs, or whether and how they use this knowledge for proof validation. This article therefore addresses the following research questions:

RQ1: To what extent do proof principles evaluated as ‘disregarded’ in a purported proof lead students to judge this purported proof as invalid (A3)?

RQ2: To what extent do students correctly identify and evaluate disregarded proof principles in invalid purported proofs (A2 & A3)?

METHOD

Sample & Items

To answer the research questions, $N = 456$ students in grades 8 to 11 from eight schools in Germany were asked to validate six purported proofs as part of an online questionnaire. All students validated the same purported proofs. The purported proofs included various mathematical content taught up to grade 8. Each of the six purported proofs was invalid as one of the three proof principles (*proof scheme*, *proof structure*,

chain of conclusions) was disregarded. Regardless of the community, none of the purported proofs could therefore be accepted as a valid mathematical proof. In four of the six purported proofs, the proof principle *proof scheme* was disregarded. The proof principle *proof structure* was disregarded in one case and the proof principle *chain of conclusions* was disregarded in another. Figure 1 shows one of the six purported proofs to be validated. In this example, the proof principle *proof scheme* was disregarded, as invalid arguments were used.

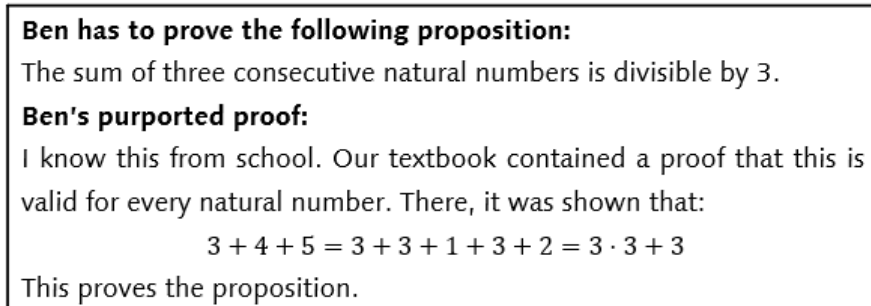


Figure 1: One of the six purported proofs to be validated by the students.

For each purported proof, the students had to evaluate the same statement for each of the three proof principles:

- *Proof scheme*: “The purported proof only uses valid arguments.”
- *Proof structure*: “The purported proof assumes what is to be shown.”
- *Chain of conclusions*: “In the purported proof, the steps build on each other.”

In this way, students evaluated whether each proof principle was fulfilled in the purported proof or not. In addition, the students had to decide whether the purported proof was a valid proof or not. They could also choose the option ‘I don’t know’. The correct answer for the purported proof in Figure 1 is that the proof principle *proof scheme* was disregarded and the purported proof is thus invalid. Due to the formulation of the statement regarding the proof principle *proof structure*, it is also correct to reject this statement for this purported proof.

Analyses

For each purported proof, regardless of whether the evaluation was correct, the number of times a student stated that a proof principle was disregarded was scored. This resulted in a score for each purported proof ranging from 0 (all three listed proof principles are fulfilled) to 3 (all three listed proof principles are disregarded). Focusing on RQ1, we first analyzed how many students rated 0, 1, 2, or 3 of the listed proof principles as ‘disregarded’ and judged a purported proof as valid or invalid. This provides insight into whether students know that proof principles evaluated as ‘disregarded’ for logical reasons result in invalid proofs (A3). The analysis was conducted in the same way for the students who stated that they do not know whether the purported proof is valid. For a more detailed analysis of the data, for all purported proofs, it was analyzed (i) which proof principle was evaluated as ‘disregarded’ and how often and (ii) how many students rated the purported proof as invalid if the proof

principle was rated as ‘disregarded’. In this way, the number of students who correctly evaluated the disregarded proof principle was determined (RQ2).

RESULTS

The analysis of all six purported proofs show that many students stated that one, two, or three proof principles were disregarded in the given purported proofs but still judged the purported proof as a valid mathematical proof. Conversely, for all purported proofs, some students stated that all three listed proof principles were fulfilled but still judged the purported proof as invalid (which might be possible). For all purported proofs to be judged, the largest group of students consisted of those who stated that they do not know whether the proof is a valid. Figure 2 provides an example of the results for one of the purported proofs to be validated.

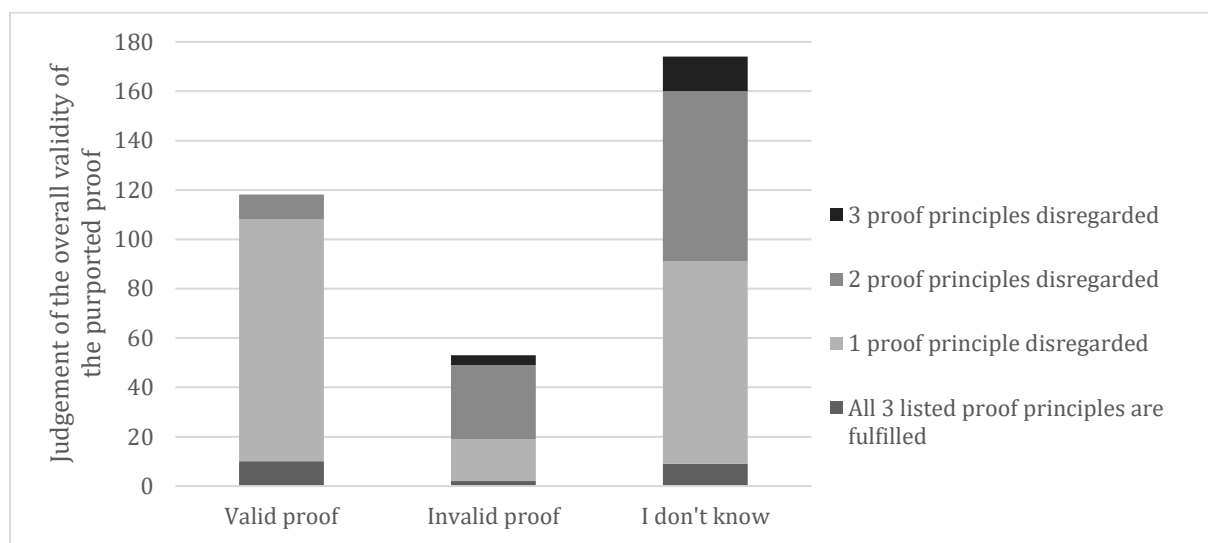


Figure 2: Descriptive data of the judgement of one purported proof, considering statements about the proof principles.

Table 1 shows the number N of students who evaluated the three proof principles as ‘disregarded’ (both in absolute and relative terms) for each of the six purported proofs. In addition, the number \tilde{N} of students that evaluated a proof principle as ‘disregarded’ for a purported proof *and* also judged the purported proof as ‘not valid’ is presented. For example, of the 279 (61.2%) students who evaluated the proof principle *proof structure* as ‘disregarded’ in purported proof 1, 24.0% judged the purported proof as invalid.

The percentage of students who correctly identified the disregarded proof principle in the purported proof and of those who further concluded the correct judgement of its validity is highlighted in grey. In purported proof 6, in which the proof principle *chain of conclusions* was disregarded, 28.2% of the students correctly identified this. Of these students, 29.7% correctly judged the purported proof as invalid. The evaluation of the proof principles shows no pattern. It is noticeable that only 7.4% of the students who correctly recognize the proof principle *proof structure* in purported proof 5 also rate the proof as invalid.

	<i>Proof scheme</i> evaluated as ‘disregarded’		<i>Proof structure</i> evaluated as ‘disregarded’		<i>Chain of conclusions</i> evaluated as ‘disregarded’	
	<i>N</i> (%)	\tilde{N} ($\tilde{\%}$)	<i>N</i> (%)	\tilde{N} ($\tilde{\%}$)	<i>N</i> (%)	\tilde{N} ($\tilde{\%}$)
purp. proof 1 (disregarded <i>proof scheme</i>)	175 (38.4%)	65 (37.1%)	279 (61.2%)	67 (24.0%)	157 (34.4%)	53 (33.8%)
purp. proof 2 (disregarded <i>proof scheme</i>)	236 (51.8%)	114 (48.3%)	200 (43.9%)	61 (30.5%)	225 (49.3%)	109 (48.4%)
purp. proof 3 (disregarded <i>proof scheme</i>)	154 (33.8%)	42 (27.3%)	210 (46.1%)	34 (16.2%)	140 (30.7%)	42 (30.0%)
purp. proof 4 (disregarded <i>proof scheme</i>)	152 (33.3%)	51 (33.6%)	215 (47.1%)	29 (13.5%)	135 (29.6%)	44 (32.6%)
purp. proof 5 (disregarded <i>proof structure</i>)	129 (28.3%)	29 (22.5%)	229 (50.2%)	17 (7.4%)	131 (28.7%)	26 (19.8%)
purp. proof 6 (disregarded <i>chain of conclusions</i>)	141 (30.9%)	35 (24.8%)	217 (47.6%)	18 (8.3%)	128 (28.1%)	38 (29.7%)

Note: \tilde{N} represents the number of students who evaluated the proof principles as ‘disregarded’ and judged the proof as ‘not valid’. $\tilde{\%}$ is the percentage of students who judged the purported proof as ‘not valid’ among those who evaluated the proof principle as ‘disregarded’ (N) (conditional probability).

Table 1: Descriptive results regarding the evaluation of the proof principles for each purported proof.

DISCUSSION

This study examined to what extent secondary school students know that disregarding proof principles in a purported proof logically leads to its invalidity. The empirical study showed that many students identified the disregarded proof principles in purported proofs but still stated that it was a valid proof. This suggests that the learners do not know that disregarding (one of) the proof principle must logically lead to the invalidity of a purported proof. Students seem to have no adequate knowledge about the essential role of proof principles for the validity of mathematical proofs (A3). The high number of students who stated that they do not know whether the purported proofs are valid could also indicate a lack of adequate knowledge about proof principles, making it difficult for them to decide on validity based on the evaluation of these proof principles (RQ1). This inadequate knowledge about proof principles could explain uncertainties in the proof (e.g., Alcock & Weber, 2005). The results in Table 1 suggest

that when the proof principle *proof scheme* is rated as ‘disregarded’, students are most likely to judge the purported proof as invalid. Although the descriptive results are still low, students probably have the best knowledge of the meaning of this proof principle as a minimum criterion. The fact that some students correctly judge an invalid purported proof as invalid, even when they state that all listed proof principles are fulfilled, suggests that they used criteria other than the three proof principles to be evaluated for proof validation (Sommerhoff & Ufer, 2019).

The results in Table 1 show that, for all six purported proofs, it is difficult for many students to identify the disregarded proof principle correctly. These results strengthen the findings that students show difficulties in correctly identifying the proof principles in the purported proofs (e.g., Alcock & Weber, 2005; Healy & Hoyles, 1998) (RQ2). For none of the three proof principles did the students consistently and correctly conclude that a purported proof cannot be valid if the evaluation of this proof principle is ‘disregarded’. Thus, these students do not yet appear to be able to use their existing knowledge about proof principles to validate purported proofs, as they seem to lack knowledge about proof principles as a minimal set of criteria.

Overall, the results suggest that students lack adequate knowledge of the logical consequences of disregarding proof principles in a purported proof. Regardless of how successfully students identify and evaluate proof principles, this makes it difficult for them to adequately use their knowledge about proof principles for validating proofs. Thus, the validation of purported proof and the criteria that can be used for this activity should have a stronger focus in mathematics education. The proof principles should not only be discussed and practiced using examples (Stylianides & Stylianides, 2009), but it should also be emphasized that proof principles must be regarded as minimum criteria and that disregarding them logically leads to the invalidity of a purported proof.

Limitation & Outlook

One limitation of this study is that the used proof principles can, so far, only be assumed to be minimum criteria for the validity of a mathematical proof. Additionally, whether students were confused by the fact that all six purported proofs were invalid should be investigated in further studies. The study does not provide any information on additional criteria that the students may have used to judge the validity of the purported proofs – especially in cases where none of the proof principles is rated as ‘disregarded’ but the overall proof is rated as ‘not valid’. Furthermore, our current analysis for Table 1 of the three proof principles in the six purported proofs did not consider whether only one proof principle was evaluated as ‘disregarded’ or whether there were several. It would be valuable to gather information on how students weight the evaluated proof principles in their proof validation. In this way, as an extension to our results, it would be possible to determine to what extent the evaluation of the proof principles is actually used in the overall judgement of the proof’s validity.

REFERENCES

- Alcock, L., & Weber, K. (2005). Proof validation in real analysis: Inferring and checking warrants. *The Journal of Mathematical Behavior*, 24(2), 125-134.
- Hanna, G., & Jahnke, H. N. (1996). Proof and proving. In A. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International Handbook of Mathematics Education* (pp. 877-908). Springer.
- Healy, L., & Hoyles, C. (1998). *Justifying and proving in school mathematics: Technical report on the nationwide survey*. Institute of Education, University of London.
- Heinze, A., & Reiss, K. (2003). Reasoning and Proof: Methodological Knowledge as a Component of Proof Competence. In M. A. Mariotti (Ed.), *Proceedings of the Third Conference of the European Society for Research in Mathematics Education* (Vol. 4, pp. 1-10). Bellaria, Italy.
- Komatsu, K., Jones, K., Ikeda, T., & Narazaki, A. (2017). Proof validation and modification in secondary school geometry. *The Journal of Mathematical Behavior*, 47, 1-15.
- Selden, A., & Selden, J. (2003). Validations of Proofs Considered as Texts: Can Undergraduates Tell Whether an Argument Proves a Theorem? *Journal for Research in Mathematics Education*, 34(1), 4-36.
- Selden, A., & Selden, J. (2015). Validation of proofs as a type of reading and sense-making. In K. Beswick, T. Muir, & J. Wells (Eds.), *Proceedings of the 39th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 145-152). Hobart, Australia: PME.
- Selden, J., & Selden, A. (1995). Unpacking the logic of mathematical statements. *Educational Studies in Mathematics*, 29(2), 123-151.
- Sommerhoff, D., & Ufer, S. (2019). Acceptance criteria for validating mathematical proofs used by school students, university students, and mathematicians in the context of teaching. *ZDM Mathematics Education*, 51(5), 717-730.
- Sporn, F., Sommerhoff, D., & Heinze, A. (2021). Beginning university mathematics students' proof understanding. In M. Inprasitha, N. Changsri, & N. Boonsena (Eds.), *Proceedings of the 44th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 102-110). Thailand: PME.
- Stylianides, A. J. (2007). Proof and Proving in School Mathematics. *Journal for Research in Mathematics Education*, 38(3), 289-321.
- Stylianides, G. J., & Stylianides, A. J. (2009). Facilitating the transition from empirical arguments to proof. *Journal for Research in Mathematics Education*, 40(3), 314-352.
- Stylianides, G. J., Stylianides, A. J., & Weber, K. (2017). Research on the Teaching and Learning of Proof: Taking Stock and Moving Forward. In J. Cai (Ed.), *First Compendium for research in mathematics education* (pp. 237-266). Reston, VA: NCTM.
- Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education*, 27(4), 458-477.

THE CULTURAL POLITICS OF KOREAN MATHEMATICS CURRICULUM REFORMS THROUGH ASIA AS METHOD

Emily Sum

Faculty of Education, The University of Melbourne, Australia

Using “Asia as Method” as an analytical framework to develop a more holistic understanding of mathematics education in South Korea between 1945 and 1988. It serves to highlight curriculum reforms by acknowledging the impact of colonization, American-imperialism and the cold-war. Also to enable a renewed knowing of how the state worked through curriculum policies to impose their ideologies of nationalism and developmentalism. It teases out the historical transformations for inter-referencing so that differences between Asian and the West can be properly explained in the future.

INTRODUCTION

Korean education has been historically and ideologically continuously influenced by American epistemologies (Kim, 2005; Lee, et al., 1988). Curriculum studies is seen as the translation and application of US/Western theories to Korean language and educational context. It has been “colonized as a secondary space wherein US curriculum conceptions are reproduced and represented” (Kim, 2005, p. 59). However, there are various sociocultural and historical aspects which made Korean schooling not at all like that of the US. So, how can we engage in curriculum inquiry solely using Anglo-US-Western theories and epistemologies in the context of a Confucian society? The diversity of Asian rich cultural heritage, ancient wisdom, historical traditions and experiences provides “alternative horizons and perspectives” (Chen, 2010, p. 212) in knowledge production. As a transnational researcher, I am interested in *the cultural politics of mathematics education*, that is “to understand the constitution of the practices of mathematics education as part of a larger cultural space where the meanings of mathematics in relations to education are constantly negotiated” (Velero, 2018, p. 107). From 1945 to 1988, various authoritarian and/or military governments, themselves backed by the US, exploited curriculum reforms to legitimize their aspirations for modernization. Thus, my questions are: What were the ideologies underpinning the curriculum reforms? And how were they being reflected in the mathematics curricula? This study focuses on reconsidering and re-examining this history of military dictatorships that maintained stability, developed economy, and controlled dissent with martial law. To posit the focus of the politics in mathematics education to show how it functioned along with the dominant ideologies within the society based on Chen’s (2010) “Asia as method”. This also contributes to the understanding of how mathematics has become an important school subject and central in the process of modernising Korean society.

ANALYTICAL FRAMEWORK & LITERATURE REVIEW

Taking Asia as method is “to multiply frames of reference in our subjectivity and worldview, so that our anxiety over the West can be diluted and productive and critical work can move forward” (Chen, 2010, p. 223). It offers “a counterpoint to normative Western theory” that “lays the groundwork for alternative ways of knowing and being”, and “challenges the geopolitics of knowledge production that creates hierarchies of knowledge often prioritizing Western scholarship” (Kester, 2023, p. 182). The postcolonial discourses can be found in the field of Asian curriculum studies (e.g. Lim & Apple, 2018; Lin, 2012) and South Korea (Kim, 2005; Kim, 2024; Lee, 2019; Rhee, 2013). US curriculum theories have been recognised as the legitimate and authorised means of Korean educational knowledge system (Lee, et al., 1988). Many Koreans have deep-rooted mindset about the superiority of the US/West, and unconscious desire to emulate their theories and ideologies. Research on mathematics curriculum beyond Western centrality has been scarce, compared to national curriculum studies (Kim, 2024; So et al., 2011). Critical curriculum theorists like Apple (1990) believes official knowledge is inherently ideological and political; and “the politics in Asia remains in a state of transition – because of this – curriculum reforms in these places involve considerable ideological work” (Lim and Apple, 2018, p. 139). Yet, the disguised ideological function of the mathematics curriculum would be impossible to apprehend without the foreground analyses that take the complex colonial, imperial, and cold-war consciousnesses in which the Korean society experienced.

RESEARCH METHODOLOGY

Document analysis was conducted to yield data and organised into themes/categories through content analysis (Bowen, 2009). National curriculum documents including vision statement, courses of study, mathematics learning standards and instructional guidelines, were the primary source used. Firstly, because “school curriculum encapsulates political aspirations for who the citizen and the nation should be” (Valero, 2018, p. 111). Secondly, the Ministry of Education controlled the process and outcome of the curriculum-making in South Korea (Sung, 2008). Curricula between September 1945 and February 1988 were retrieved from the National Curriculum Information Center. Initial analyses identified some keywords including “nation”, “productivity”, “usefulness” along with, “anticommunism”, “democracy”, “moral” and “ethics”. Other forms of government publications including newspapers, press releases, official reports and documents, related to educational policies during same period were retrieved from the National Archives of Korea. Data were analysed and organised into themes using Asia as method as framework that inquired into decoloniality and transformative practices from a perspective that less commonly adopted in Asian educational research (Kester, 2023). Additional analyses were conducted to reconceptualise the ideologies *colonialism*, *Americanism*, *imperialism*, *nationalism*, to describe how power unfolds in and through mathematics education.

ANALYSIS & FINDINGS

Colonialism

In the late Joseon Korea (1392-1910), both traditional and Western mathematics were taught in schools (Lee et al., 2009). During Japanese colonization (1910-45), the colonizer had implemented their own educational system, curriculum and pedagogy on the terrain of the colonized; Korean Language, History and Geography were suppressed (Lee, 2003; Ministry of Education, Science and Technology [MEST], 2012). Students were forced to learn Western mathematics in Japanese – the national language – instead of Korean (Lee, et al., 2009). The teaching hours of arithmetic for boys was more than girls in elementary schools (Kim, et al., 2013).

American-imperialism and the Cold-War

The process of de-decolonization began after the liberation of Japanese occupation but was interrupted by the US military government (1945-48) and the Korean War (1950-53). Given the country's economic status, the mission of rebuilding Korean education was clear – industrialisation (Ministry of Education [MOE], 1953). Under the political influence of the US and US-educated Rhee Syng-man regime (1948-60), teacher training was conducted by American education delegation and specialists according to US Progressive educational theory (MEST, 2012). And to justify the coup backed by the US during the cold-war era, anticommunism education, moral education and vocational education were the major principles in the first national curriculum (MOE, 1955). The mathematics curriculum and teaching models were influenced by Deweyan theory of education (Lee, 2003). Progressivism was emphasized and mathematics was practical and regarded as a tool of living, the mathematical content was in low level and unstructured (Kim, et al., 2013; Paik, 2004). There was an attempt to connect mathematics with citizenship, as documented in the curriculum guideline – “arithmetic is a knowledge that every single citizen should understand, ... it should not be something magical, mysterious, incomprehensible” (American Education Team, 1954, p. 177). Under President Park Chung-Hee and his military regime (1963-79), the second curriculum was announced in 1963. Anticommunism and moral life became a stand-alone subject in elementary and middle schools that should also be implemented across all subject areas (Government Gazette, 1963). The mathematics curriculum was designed based on Herbart's Essentialism. It was more systematic and subject centred, with emphasises on logical and theoretical aspects of mathematics (Park, 1997). The high school mathematics curriculum started to offer Common Mathematics, Mathematics I, and Mathematics II. It was differentiated where both practical mathematics and advanced mathematics courses were available for students in the second-year and third-year of high school (MOE, 1963). This structure remains the same in the current curriculum. During the development of the second curriculum, New Math movement had begun in America, so Korean mathematics educators incorporated set theory as a curriculum content of the education university for elementary school teachers (Government Gazette, 1969). The third curriculum in 1973 renamed the

subject “arithmetic” to “mathematics” in elementary school. The contents were aligned with the New Math, including a significant amount of abstract mathematics. Unlike some European countries where reforms affected mainly secondary and higher education (Gosztonyi, et al., 2023), concepts and notations of set theory was introduced in early elementary grades, along with algebraic properties as common foundation to facilitate better learning. Moreover, set theory as a way of thinking was highlighted as an objective in high school curriculum. The reform concerned with pedagogical as much as content, including explaining, visualising, and manipulating concrete materials (Kim, et al., 2013; MOE, 1973). The metal and mathematical structures that followed Piaget and Bruner’s educational psychology were incorporated hoping to support mathematics teaching and discovery learning in Korean school settings. The fourth curriculum, the first revision of the New Math, was made in 1981 under the new governance of Chun Doo-hwan (1980-88). Some contents of the New Math were removed, with an emphasis on basic skills, which resulted in a considerable retreat from the New Math (MOE, 1981). This reflected the Back-to-Basics movement in the US, emphasizing a reduction in learning contents and lowered the level of difficulties. Changes also emphasized the ability and the integrative approach to problem solving (Paik, 2004). These reforms were mainly based on the implementation of the US, without considering the complexity of post-war situation in Korea. Certain knowledge was included, while other knowledge and experiences were excluded from the official curricula, this yields a conceptual and political understanding of what counted as school mathematical knowledge to be (re)produced and transmitted. I consider this as the politics of mathematical knowledge (Frankenstein, 1998), and the choice of what considered as official knowledge to be taught in school was controlled by pro-American authorities, mathematics education was Americanized during the cold-war.

Nationalism to Developmentalism

In 1961, education in Korea entered a new epoch under the totalitarian governance of Park. The educational policy was mainly focused on creating human capitals to serve country’s economic development. It was clear that the second national curriculum emphasised on productivity and usefulness to reconstruct economy for the prosperity and development of the nation and society (MOE, 1963). All curricular and extracurricular activities should be designed for students to acquire sufficient scientific and technological capability for employment and higher education (Government Gazette, 1967). The state had been powerful in regulating the educational field to legitimate the reproduction of the ideologies and agencies. From the 1960s, the military government had been justified mathematics and science as core subjects for economic development and global competitiveness. Vocational schools were established in 1963 aimed to prepare students even with lower academic performance could become technical experts who were most needed for various growing industries (Lee et al., 2022). At the same time, high school mathematics prepared students for the rigorous knowledge of university education creating pathways for Science, Technology, Engineering, and Mathematics (STEM) and accessed to higher-paying jobs. In the

spectacular growth of capitalism in East Asia among the so-called Four Litter Tigers (Hong Kong, Singapore, South Korea, and Taiwan). The significant development in STEM education and strong educational policies supported the technological advancement, manufacturing industries in the 1970s, to electronics and automobiles in the 1980s (Seth, 2017). Many Koreans hold the assumption that “those who acquired more education and technical skills, particularly in the sciences and engineering sectors, could easily get decent jobs in big firms, which guaranteed job security, good salaries, and excellent perks” (Yang, 2018, p. 49). Industrialization and modernization would have been impossible if not for the high emphasis placed on economic success, with self-sacrificing, or sometimes self-exploiting labour (Ringmar, 2005). That was the Korean STEM education in its infancy, which helped the nation to adapt to the rapidly changing knowledge-based economy in the millennium. The governments of Rhee and Park both sought to make use of the educational system as a means of political indoctrination and control. When comparing the two high school mathematics curricula during Park’s governance, the objectives shifted from the development of an individual’s thinking and reasoning (MOE, 1963) to the contribution of national development (MOE, 1973). The educational policies clearly documented how the state politicised school mathematics with the association between personal development and nation prosperity, and to justify the learning of certain mathematics for ideological legitimacy. Moreover, economic modernization projects such as *Saemaul Undong*, exploited *Saemaul* education to train the rural population about the economic benefits of adopting agricultural and industrial technology in protecting themselves from communist neighbours (Korea Institute of Public Administration, 1980). As Chen (2010, p. 123) interprets “nationalism has played a critical mediating role in that it allows for the projection of an imaginary enemy, which in turn is used by the state to justify its suppression of the internal drive and desire for democracy”. Developmentalism and capitalism drive citizens for tangible and financial benefits, allowing miraculous economic growth in a short period of time – the Miracle of Han River. School mathematics was directed to meeting the needs of industry and business as the primary goal. Integrating mathematics education into economic modernization projects by focusing on producing and cultivating technical professionals to accelerate industrialization. The conditions for thinking and reflecting American-imperialism on education were depoliticised, postponed, and channelled into economic development. It was the transplantation of American-imperialism and capitalism; South Korea had shifted from an agricultural economy to an industrialized economy. That was the kinds of citizens and citizenship that the state wish to promote. Solidary and nationalism were the means to unite the nation and overcome the crisis, and Confucian traditions of filial piety and obedience for the bureaucracy were incorporated into school curriculum in 1973. “Nationalism took up this Confucian guise, when necessary, either to legitimize authority or to mobilize resources for political purpose” (Kim, 2017, p. 24), which promoted education to facilitate human resources and invoked patriarchal kinship patterns to create a docile labour force (Park, 1999).

Inter-referencing

The active importation of US educational philosophy, ideology and values, through pro-American political regimes, had shaped Korean mathematics education in the twentieth century. Major changes in the reforms had been dominated by those who seized the political power in achieving their ideologies, as well as political and economic objectives. As Chen (2010) explained, pro-American Asian countries such as South Korea and Taiwan had secured the state power using anticommunism as the fundamental national policy and the dominant ideology to justify authoritarianism and developmentalism. The sociohistorical fact that Korea underwent Japanese colonization, the US military government, and the cold-war as US ally, has made Taiwan an important point of reference. US policies toward Taiwan feature similar support for dictatorial regime, the Kuomintang, which developed national economy, and served American interest. “The installation of anti-communist–pro-Americanism structure in the capitalist zone of East Asia” (2010, p. 7), resulted in highly defensive forms of nationalism that prohibited to questions about American-imperialism, and still struggling for emancipation from their control. This is why Taiwanese mathematics curriculum reform is worthy to be examined from an Asian dichotomy, as it may provide another counterpoint to normative Western theory/scholarship, also alternative ways of knowing and being (Kester, 2023). It may offer a contextual, historical grounding to de-colonising and de-Americanising mathematics curriculum while considering the power relationships in the formation specific to East Asian context.

CONCLUSION

Thanks to government initiatives during nation building and state making, mathematics education had shifted its focus quickly from developing citizenship to developing economies for industrialisation, making mathematics a valuable subject in South Korea. One of the lasting legacies of the cold-war is the installation of anticommunism, democracy, and pro-Americanism structure, whose overwhelming consequences are still with us today. American was the dominant educational system of reference; the authoritarian governments had adopted mathematics curricula and practices to secure and increase their power, plus interests, in the larger social arena. American ideologies dominated the official knowledge of mathematics – the cultural capital of Koreans. The adoption of New Math was a fear of being left behind in the global competitions and catching up with other neighbouring countries. Thus, the influences of US hegemony on the mathematics education within East Asian region need to be critically re-examined so that the underpinning beliefs and thoughts can be acknowledged, and comparative studies of East Asia and the West can be conducted more meaningfully.

References

- Apple, M. W. (1990). The politics of official knowledge in the United States, *Journal of Curriculum Studies*, 22(4), 377-400.
- American Education Team (1954). *Curriculum handbook for the schools of Korea*. The United Nations Korean Reconstruction Agency and the American-Korean Foundation.

- Bowen, G. A. (2009). Document analysis as a qualitative research method, *Qualitative Research Journal*, 9(2), 27-40. <https://doi.org/10.3316/QRJ0902027>
- Chen, K.-H. (2010). *Asia as method: Towards deimperialization*. Duke University Press.
- Frankenstein, M. (1998). Reading the world with maths: Goals for a critical mathematical literacy curriculum. In P. Gates (Ed.), *Proceedings of the first international mathematics education and society conference* (pp. 180–189). Nottingham.
- Gosztonyi, K., van den Heuvel-Panhuizen, M., Makinae, N., Shimizu, S., van Zanten, M. (2023). International Co-operation and Influential Reforms. In: Shimizu, Y., Vithal, R. (eds) *Mathematics Curriculum Reforms Around the World*. Springer.
- Kester, K. (2023). Revising ‘Asia as Method’ in education research: Problems and prospects, *Asia Pacific Education Review*, 24, 181-186.
- Korea Institute of Public Administration (1980). *Curriculum materials of the Saemaul education*. Republic of Korea.
- Kim, C. Y. (2005). Post-colonialism and reconstruction of Korean curriculum studies, *Journal of Curriculum Studies*, 21(1), 57-75.
- Kim, D. (2024). Taking “Asia as method” seriously in decolonizing comparative curriculum studies in South Korea, *Journal of Curriculum Studies*, 207-219.
- Kim, K.-D. (2017). *Confucianism modernization East Asia: Critical reflections*. Palgrave Macmillan.
- Kim, J. Han, I., Park, M., & Lee, J. K. (2013). *Mathematics education in Korea: Curricula and teaching and learning practices*. World Scientific.
- Lee, J. J., Adams, D. & Cornbleth, C. (1988). Transnational transfer of curriculum knowledge: A Korean case study. *Journal of Curriculum Studies*, 20(3), 233-246.
- Lee, S.-G., Noh, J., & Song, S.-Y. (2009). Educational policy and curriculums of Korean school mathematics in the late 19th and early 20th century, *Journal of Korean Society of Mathematics Education*, 23(4), 1093-1130.
- Lee, Y. (2003). Politics and theories in the history of curricular reform in South Korea. In W. F. Pinar (Ed.), *International Handbook of Curriculum Research* (pp. 541–552). Routledge.
- Lee, D.-H., Kwak, S.-G., Kim, J.-W., Park, D.-J., Yang, J.-H., & Lee, M.-H. (2022). *Education in South Korea: Reflections on a seventy-year journey*. Palgrave Macmillan.
- Lee, Y. (2019). A critical dialogue with ‘Asia as method’: A response from Korean education, *Educational Philosophy and Theory*, 51(9), 958-969.
- Lim, L., & Apple, M. W. (2018). The politics of curriculum reforms in Asia: Inter-referencing discourses of power, culture and knowledge. *Curriculum Inquiry*, 48(2), 139-148.
- Lin, A.M.Y. (2012). Towards transformation of knowledge and subjectivity in curriculum inquiry: Insights from Chen Kuan-Hsing’s ‘Asia as Method’, *Curriculum Inquiry*, 42(1), 153-178. <https://doi.org/10.1111/j.1467-873X.2011.00571.x>
- Ministry of Education (1953). *Rebuilding education in the republic of Korea: The final report of the UNESCO, UNKRA educational planning mission to Korea*. Republic of Korea.

- Ministry of Education (1955). *National curriculum*. Republic of Korea [ROK].
- Ministry of Education (1963). *National curriculum*. ROK.
- Ministry of Education (1973). *National curriculum*. ROK.
- Ministry of Education (1981). *National curriculum*. ROK.
- Ministry of Education, Science and Technology (2012). *Analysis on development and achievement of compulsory elementary education after the Korean War*. ROK.
- Government Gazette (1963). No. 3373, February 15, 1963. ROK.
- Government Gazette (1967). No. 4624, April 15, 1963. ROK.
- Government Gazette (1969). No. 5181, February 24, 1969. ROK.
- Park, T.-K. (1999). Confucian values and contemporary economic development in Korea, In T. Brook and H.V. Luong (Eds.) *Culture and Economy: the shaping of capitalism in eastern Asia* (pp. 125–134). University of Michigan Press.
- Paik, S.-Y. (2004). Mathematics curriculum in Korea. Paper presented at ICME10, Copenhagen, Demark. <http://matrix.skku.ac.kr/For-ICME-11/ICME/cp2.pdf>
- Park, K. M. (1997). School mathematics curriculum in Korea, *Journal of Korea Society of Mathematics Education*, 1(1), 43-59.
- Rhee, J.-E. (2013). Methodology of leaving America for Asia: Reading South Korea's social studies textbooks through Chen Kuan-Hsing's Asia as Method. *Qualitative Research in Education*, 2(3), 328-354. <https://doi.org/10.4471/qre.2013.31>
- Ringmar, E. (2005). *Surviving capitalism: How we learned to live with the market and remained almost human*. Anthem Press.
- Seth, M (2017). South Korea's economic development, 1948-1996. <https://doi.org/10.1093/acrefore/9780190277727.013.271>
- So, K., Kim, J., & Lee, S. (2012). The formation of the South Korean identity through national curriculum in the South Korean historical context: Conflicts and challenges, *International Journal of Educational Development*, 32(6), 797–804.
- Sung, Y. K. (2008). The characteristics of national curriculum decision-making in Korea: Discovering metaphors, *The Korean Journal of Educational Methodology Studies*, 20(2), 41–56.
- Valero, P. (2018). Human capitals: School mathematics and the making of the homus Oeconomicus. *Journal of Urban Mathematics*, 11(1&2), 103-117. <http://education.gsu.edu/JUME>
- Yang, M. (2018). *From miracle to mirage. The making and unmaking of the Korean middle class, 1960 – 2015*. Cornell University Press.

EXAMINING EQUITABLE PARTICIPATION WITH K-12 TEACHER LEADERS

Eva Thanheiser, Kathryn Roman, Ruth Heaton
Portland State University

Many organizations and researchers in mathematics education call for antibias practices, but addressing hidden biases remains challenging. Equity QUantified In Participation (EQUIP), a quantitative tool designed to analyze classroom participation patterns by social markers, helps teachers identify these subtle inequities. This paper describes a professional development (PD) intervention introducing EQUIP to five teacher leaders (TLs) and three coaches in a predominantly Latine district, focusing on setting up EQUIP to identify and address participation inequities in their classrooms. The TLs implemented EQUIP in their classroom and gathered participation data twice. Findings revealed that TLs took up EQUIP in different ways and became more aware of hidden participation patterns. TLs used what they learned to adjust their practice. These changes resulted in more balanced participation and improved student engagement, particularly for students traditionally marginalized in mathematics classrooms.

RATIONALE OF THE STUDY AND RESEARCH QUESTIONS

Many organizations in mathematics education as well as researchers are calling for antibias mathematics education (Thanheiser, 2023; TODOS: Mathematics For All, 2020; Wilkerson & Berry III, 2020). The Association of Mathematics Teacher Educators (AMTE), in particular, urges teachers to “actively work to be anti-racist in our acts of teaching, research, and service” (Steele & Burton, 2020 para 5). Similarly, TODOS (2020) asserts that an antiracist position in mathematics education is a pledge to dismantle systems and structures that maintain racism within teaching and learning mathematics from challenging belief systems that perpetuate microaggressions to disrupting the role mathematics classes play in pushing students out of schooling p. 2).

To achieve this, we need to recognize the implicit biases we (as teachers) hold in the math classroom and how these biases impact *who* is participating. Equity QUantified In Participation (EQUIP) (Reinholz & Shah, 2018) is a quantitative tool designed to analyse classroom participation patterns by social markers and helps teachers identify these subtle inequities.

In response to this pressing need, we introduce a professional development (PD) focussed on introducing the use of EQUIP in one’s own classroom. Our **research questions** are:

- How can EQUIP be introduced to teachers so they can use it in their own classrooms?

- What did teachers learn by using EQUIP in their own classrooms?
- What changes did teachers make based on what they learned about their own classrooms?
- How did those changes affect classroom participation?

THEORETICAL FRAMEWORK AND LITERATURE

EQUIP is a quantitative tool that allows teachers/researchers to track student participation and their discourse patterns in the classroom by social markers, such as ethnicity and gender (Reinholz & Shah, 2018). Sometimes educators use the EQUIP tool as designed (e.g., Byun et al., 2023), others have opted to use tallies to track participation (e.g., Herbel-Eisenmann & Shah, 2019). Regardless of the method used, educators gain insights into the subtle inequities present in their classes. For example, researchers have found men contributing more than women (Ernest et al., 2019), white boys answering why questions disproportionately (Reinholz & Shah, 2018), Latine students being underrepresented in participation (Reinholz & Shah, 2018; Reinholz & Wilhelm, 2022), and certain students dominating discussions (Byun et al., 2023; Reinholz et al., 2022).

Although the participation patterns tracked by EQUIP are not new or surprising in mathematics classrooms, how teachers changed these participation patterns once they were aware of them in their own classrooms is notable. For example, preservice teachers (PSTs) in a methods class suggested using equity sticks (Byun et al., 2023) and teacher-researchers in Herbel-Eisenmann and Shah's (2019) study planned to be more intentional about who they called on, by encouraging students outside of class to participate more and changing students' seats to support more equitable participation. As these examples show, EQUIP can help teachers see the inequities inside their classrooms and assess how instructional changes affect these inequities. However, it has been noted that using EQUIP can be time consuming (Byun et al., 2023; Cunningham et al., 2024) as it requires videotaping and careful coding of lessons. Few studies have documented teachers using EQUIP on their own classroom data (exceptions include Block & Males, 2020; Byun et al., 2023; Herbel-Eisenmann & Shah, 2019). Even fewer focussed on how to set up EQUIP (exceptions include Cunningham et al., 2024). Therefore, the purpose of this paper is to collaboratively examine the use of EQUIP with K-12 teacher leaders (TL), and thus contributing to the understanding of how this tool can support equitable participation in mathematics education.

METHODOLOGY

We developed and implemented a professional development (PD) that introduced the use of EQUIP. The PD sessions were videotaped, transcribed, and then analyzed with the following questions in mind: How did the TLs take up EQUIP? What did the TLs learn? What goals did they set for their teaching? Participants were five K-12 TLs, two elementary, two middle, and one high-school teachers, and 3 coaches from a primarily

Latine district. The EQUIP PD was part of a larger anti-bias PD effort, however, in this paper we focus only the EQUIP part. Below we lay out our activities.

PD Activities

PD Session 1: Introduction of the EQUIP Tool to the TLs Before introducing EQUIP, the TLs collaboratively engaged in a mathematics activity. The video of the math activity was analysed using EQUIP to track TL participation. (This was shared with the TLs in the next PD session). Then EQUIP was introduced as a tool that can be implemented regardless of pedagogy and will allow the teacher to better understand participation patterns. The focus on student participation (not pedagogy) was emphasized throughout. The TLs each created a general EQUIP account and were encouraged to explore the existing tools. Once TLs were familiar with the general EQUIP setup, we shared a predesigned classroom setup in EQUIP which allowed the TLs to collaboratively code a video of that pre-designed classroom. We played small increments of the video and then discussed how to code it. Discussions were analysed for instructor activity and teacher questions to identify potential struggles with the tool. After coding some of the video together we shared the complete EQUIP coding of the video with the teachers for discussion. Discussions were analysed to identify what TLs noticed. For example, one TL noticed how some of the analysis tools allow one to change between demographic categories to view results like by gender or ethnicity. Another TL observed that boys interrupted more frequently than girls (we note the limitation of the gender binary of this data). Next, each TL set up their own classroom (using their student rosters and categories they wanted to explore) in order to implement EQUIP in their own classroom. TL were asked to record 10 minutes of their own classroom and use EQUIP to analyse it before the next PD session.

Between PD Sessions 1 and 2. TLs implemented and analysed EQUIP in their own classrooms. TLs participated in interviews by school level groups.

PD Session 2: Share Experiences of Implementing EQUIP in Classrooms. TL shared and discussed what they learned from their implementation of EQUIP. The results of using EQUIP on the math activity of the prior PD was shared with the TLs. The TLs discussed what they noticed in the results. TLs were asked to implement EQUIP again before the next PD session.

Between PD Sessions 2 and 3. Teachers implemented and analysed EQUIP in their own classrooms. Teachers participated in interviews by school level groups.

PD Session 3: Share Experiences of Second Implementation of EQUIP in Classrooms. TLs shared and discussed what they learned from their overall implementation of EQUIP.

Data Sources and Analysis

The PD and interviews were videotaped, transcribed, analysed using three questions:

- What was the implementation of EQUIP like?
- What Did TLs Learn?
- What goals did they set for their teaching?

The Math part of the PD was analysed using EQUIP (to be shared at the next PD session). While the teachers videotaped their own classroom and used EQUIP to code it, we did not use that data in our analysis. Our analysis focuses on teacher participation in the PD and reflections on using EQUIP.

Sample analysis of an interview segment. Interview response of the kindergarten TL, Reina (all names are pseudonyms), after the first implementation of EQUIP. For this TL, the coach, Catalina, came in to do the video recording and supported the coding.

Transcript	Coding
<p>Interviewer:</p> <p>What was it like using it [EQUIP]?</p> <p>And then what you learned from it?</p>	<p>Prompting for</p> <ul style="list-style-type: none"> • Engagement with EQUIP • Learning from EQUIP • Setting goals for teaching based on using EQUIP
<p>Reina: Okay, so videoing was not not difficult at all. Catalina [coach] did it all. ... As we were videoing, I really wasn't thinking about the video, because, what was, I was just concentrating on the lesson itself. ...</p>	<p>Engagement with EQUIP via videotaping and then coding the videotaped lesson. The engagement was easy because the coach came in and did the videotaping so Reina did not have to worry about it.</p>
<p>Reina: ... It was quite revealing. It was very revealing, but as we were also using EQUIP [to code] we noticed some very interesting things. ... I noticed several, a couple of students just dominating even as we were asking them, you know, not to. It was quite difficult for some of the kids. They are 5, after all, but they, it was very difficult for certain students not to share or not to respond, even though we</p>	<p>Learning from EQUIP</p> <p>Some students dominated the discussion even though they were asked not to.</p> <p>Even though Reina used the “equity sticks” to call on people, the same students still dominated the discussion.</p>

<p>were using name sticks as well, where I was calling out kids.</p>	
<p>Reina: It was just really hard for them. So it's trying to get as much, as much as the, the involvement as possible from a range of students, but I noticed that there were a couple that were just really, yes, very difficult, but call out and such.</p>	<p>Setting goals:</p> <p>Goal is to get a wider range of students to participate</p>

RESULTS AND INTERPRETATION

How Did the TLs Take Up EQUIP?

After the introduction to EQUIP the TLs took it up in different ways during the first implementation. The two elementary teachers recorded videos and analysed them with EQUIP with support of a math coach. One middle school TL tracked students' discourse during instruction by tallying participation, the other middle school teacher was not able to implement in the first round. The high school teacher used the ideas of EQUIP to reflect on a previously recorded video of her teaching. These implementations mirror the literature with the exception of the high school teacher. In the second implementation all TLs videotaped their classroom and coded the video using EQUIP with support of either a coach or a graduate student.

What Did TL Learn In The First Implementation? And What Changes Did They Make Based on That Learning?

Both elementary school teachers recognized how one particular student talked more than the other students in each of their classes. "I noticed how much more...certain little people were dominating the conversation over everyone else [from] really watching the video." While the teachers seemed to be aware of these students' domination in the classroom the extent of it was explicated through EQUIP. "He's not really letting anyone else answer. I knew that, but I really didn't know how much." The kindergarten TL decided to be more intentional about managing participation when using name sticks, becoming "a little more manipulative, with those name sticks." She also talked to students beforehand about their participation. For example, she explained how she told a student, "We need, you need to respect everyone else's learning as well. I know you'd like to talk too and we can talk at a different time. But during this you need to let everyone else talk as well." Another thing they noticed was the different lengths of student contributions. Most of their students contributed with 1-4 words rather than longer sentences. As a result, the TLs started being more

purposeful in distributing opportunities to participate across the class and using sentence frames to encourage longer student responses.

The middle school TL, who implemented EQUIP in the form of tallying, found that newcomers and students with IEPs participated less frequently than he had realized. In accordance with his noticing, he shifted students' seating: "They are literally in the first two rows now, and they're seated closer to the podium as well, so I have quicker access to them." He reported changes in participation patterns.

The high school TL recognized that familiarity amongst the students impacted how they collaborated with each other in small groups. She explained, "The overall theme was like the chemistry of the team members themselves, if they were more talkative, then they would kind of bounce off each other. And then if they're both silent, it would never happen." As a result, she decided to change groups more often "and then build in at least a little bit of icebreaking time when they are in their new groups." She noted how "it seemed like it's gotten better with that."

In summary, all teachers who implemented EQUIP learned something new about their teaching and their students' participation through engaging in various forms of EQUIP-ing.

What Did TLs Learn In the Second Implementation? How Did The Changes Made Affect Classroom Participation?

The elementary TLs recognized that their interventions of redistributing opportunities to participate as well as implementing sentence frames was successful. The elementary school coach attributed these improvements to their use of EQUIP. She explained, "I think EQUIP helped highlight [the issue] and they [the elementary school teachers] just continued to work on having kids speak in a complete sentence." As a result, she noted, "the kids know that [a response is] not one word."

The middle school teachers learned about gender differences in participation in their classroom. One teacher noticed that the same boys frequently called out more than girls: "Gender is going to be almost all boys...Because a lot of the girls in that class do not yell out. It's because it's the same boys that yell out the whole period." Although she was already aware of this, EQUIP helped her recognize *how often* it occurred: "It was surprising to see how many times I was interrupted by the same student, like, I knew it was a lot...but...it's so disruptive." In contrast, the other middle school teacher observed that girls spoke more often than boys, specifically noting, "The only surprise is [student name]. She'll generally whisper responses. And that's what I caught on the video." Thus, the teacher was able to determine that the student, despite whispering her responses, was actively participating in class discussions.

The high school teacher recognized different ways of participating and how the chemistry among small groups still influenced student interaction. She noted how writing things down counted as participation: "I think probably...seeing the one or two

students writing things down...was not a surprise, but interesting realizing oh yeah, that's...still considered as student discourse." She also let students choose their own seats for the first time in five years and observed, "There was one group of boys that didn't really talk at all. But again, they were kind of like, thrown together a little bit rather than, like, didn't choose to sit together." She added, "I am very careful with my seating charts, usually."

Summary

This short intervention allowed teachers to uncover often hidden participation patterns. During the first implementation of EQUIP, teachers took different approaches: elementary teachers analyzed video recordings, one middle school teacher tallied participation, and the high school teacher reflected on a prior video. Even though they took different approaches, some requiring much less time commitment than others, they all learned. As such, one potential contribution of this work is that EQUIP can be implemented in various ways in K-12 settings to allow TLs to learn about the participation patterns in their classrooms. They discovered the impact of dominant speakers, varying contribution lengths of their students, and the impact of seating arrangements on participation. In the second implementation, all teachers used EQUIP to video record and code their classrooms. They observed improvements in redistributing participation and longer contributions, and noticed gender differences in participation patterns. The high school teacher also reflected on how group chemistry contributed to student discourse and writing counted as participation. Overall, EQUIP helped all teachers gain new insights into student participation and their teaching practices. In addition, as little as two implementations can allow teachers to recognize the impact that changes in their teaching practices make on student participation patterns.

RELEVANCE TO A PME AUDIENCE

Importance, Value, or Interest to the Community

Our proposal focuses on how coding one's own video data using EQUIP can help teachers recognize and address subtle inequities in their classrooms. This work is valuable to the mathematics education community because it offers a practical, research-backed way to improve equity in classroom participation. By sharing various ways of how EQUIP can be implemented, we hope to support teachers in creating more inclusive and equitable learning environments—something that is crucial in today's classrooms.

REFERENCES

- Block, S., & Males, L. M. (2020). Examining equitable participation in a professional learning community using the EQUIP tool. In W. M. Smith & L. C. Augustyn (Eds.), *Proceedings of the ninth annual Mathematics Teacher Education Partnership (virtual) conference* (pp. 43–46). Washington DC: Association of Public and Land-grant Universities.

- Byun, S., Shah, N., & Reinholz, D. L. (2023). When only white students talk: EQUIP-ing prospective teachers to notice inequitable participation. *Mathematics Teacher Educator*, 11(3), 155–168. <https://doi.org/10.5951/MTE.2022-0018>
- Cunningham, J., Takahashi, S., & Hirschboeck, K. (2024). *Using the EQUIP observation tool to advance equity in math classrooms* (pp. 1–12). WestEd. <https://www.wested.org/resources/using-the-equip-observation-tool-to-advance-equity-in-math-classrooms/>
- Ernest, J. B., Reinholz, D. L., & Shah, N. (2019). Hidden competence: Women’s mathematical participation in public and private classroom spaces. *Educational Studies in Mathematics*, 102(2), 153–172. <https://doi.org/10.1007/s10649-019-09910-w>
- Herbel-Eisenmann, B., & Shah, N. (2019). Detecting and reducing bias in questioning patterns. *Mathematics Teaching in the Middle School*, 24(5), 282–289. <https://doi.org/10.5951/mathteacmiddscho.24.5.0282>
- Reinholz, D. L., Johnson, E., Andrews-Larson, C., Stone-Johnstone, A., Smith, J., Mullins, B., Fortune, N., Keene, K., & Shah, N. (2022). When active learning is inequitable: Women’s participation predicts gender inequities in mathematical performance. *Journal for Research in Mathematics Education*, 53(3), 204–226. <https://doi.org/10.5951/jresmetheduc-2020-0143>
- Reinholz, D. L., & Shah, N. (2018). Equity Analytics: A Methodological Approach for Quantifying Participation Patterns in Mathematics Classroom Discourse. *Journal for Research in Mathematics Education*, 49(2), 140–177. <https://doi.org/10.5951/jresmetheduc.49.2.0140>
- Reinholz, D. L., & Wilhelm, A. G. (2022). Race-gender d/discourses in mathematics education: (Re-)producing inequitable participation patterns across a diverse, instructionally-advanced urban district. *Urban Education*, 0(0), 1–31. <https://doi.org/10.1177/00420859221107614>
- Steele, M., & Burton, M. (2020). AMTE Statement on Systemic Racism [Press Release]. <https://amte.net/civicism/mailling/view?reset=1&id=128>
- Take a Test. (n.d.). Project Implicit. Retrieved October 4, 2024, from <https://implicit.harvard.edu/implicit/takeatest.html>
- Thanheiser, E. (2023). What is the mathematics in mathematics education? *The Journal of Mathematical Behavior*, 70, 101033.
- TODOS: Mathematics For All. (2020). The mo(ve)ment to prioritize antiracist mathematics: Planning for this and every school year. <https://www.todos-math.org/assets/The%20Movement%20to%20Prioritize%20Antiracist%20Mathematics%20Ed%20by%20TODOS%20June%202020.edited.pdf>
- Wilkerson, T. L., & Berry III, R. Q. (2020). A statement on George Floyd, Breonna Taylor, and Ahmaud Arbery [Press Release]. <https://www.nctm.org/News-and-Calendar/News/NCTM-News-Releases/A-Statement-on-George-Floyd,-Breonna-Taylor,-and-Ahmaud-Arbery/>
-

ANALYSIS OF STUDENTS SELF-WRITTEN WORD PROBLEMS RELATED TO ARRANGEMENT WITHOUT REPETITION

Charlott Thomas and Birte Pöhler

University of Potsdam, Germany

Students often struggle with combinatorial problems, primarily due to a lack of conceptual understanding. Unlike studies focusing on procedural knowledge, this research investigates 17- to 18-year-olds' conceptual understanding of arrangements without repetition. A test with 161 students required the formulation of a text task for a given equation, revealing substantial challenges and gaps in understanding. These findings highlight the need for deeper exploration of conceptual understanding, as combinatorics fosters rich mathematical thinking and is integral to the German curriculum.

INTRODUCTION

Research indicates that combinatorial tasks can be particularly challenging for students in secondary education (e.g., Lamanna et al., 2023). Studies on challenges in combinatorial tasks is divided into three main areas (Höveler, 2014): task characteristics, personal factors, and teaching variables. However, this study focuses on task characteristics, specifically as students formulate word problems from equations, as the focus is mainly on the combinatorial task and the students' solution behaviour is analysed. Studies on task characteristics cover combinatorial operations (permutations, arrangements, and combinations, with or without repetition) (Batanero et al., 1997a; Fischbein & Gazit, 1988; Lamanna et al., 2022), implicit combinatorial models (selection, partition, distribution) (Batanero et al., 1997b, Lamanna et al., 2022), and the type and number of elements combined (Fischbein & Gazit, 1988). Research shows varied success rates based on task types and student age. For instance, Fischbein and Gazit (1988) found that permutation tasks were most challenging, while Batanero et al. (1997a) observed that older students succeeded more with permutation tasks. Selection-based tasks generally saw higher success rates (Batanero et al., 1997b; Lamanna et al., 2022), possibly due to the selection model's prevalence in textbooks. Finally, students tend to perform better when combining simpler elements, such as numbers or letters, than more complex objects (Batanero et al., 1997a; Fischbein & Gazit, 1988).

A significant reason for these difficulties is a lack of conceptual understanding of combinatorial concepts and operations (e.g., Hefendehl-Hebeker & Törner, 1984). In contrast to procedural knowledge, conceptual understanding goes beyond mere memorisation and enables students to develop a deeper understanding of a topic or concept. Furthermore, this understanding enables students to apply learnt concepts to new situations and problems (Hiebert & Carpenter, 1992). In order to gain insight into

the students' conceptual understanding, it is possible to alternate between verbal, symbolic and graphic representations (e.g., Kilpatrick et al., 2001). This study starts at this point and investigates students' conceptual understanding of combinatorics at the end of their school career by analyzing self-written word problems within Lockwood's (2013) model of combinatorial thinking. It focuses on arrangement without repetition, the first combinatorial operation taught per German curricula (KMK).

THEORETICAL BACKGROUND

In addition to Halani's (2013) model, which describes three approaches (thinking in union, deletion and equivalence classes) that can help learners develop a plan for solving combinatorial problems, *the model of students' combinatorial thinking* focuses on the components *formulas/expressions*, *counting processes* and *sets of outcomes* (Lockwood, 2013). She emphasised that students may present their solutions to combinatorial problems using one or more components and their errors may be related to one or more components or the lack of correct connections between these components (Lockwood, 2013). *Formulas/expressions* are defined as mathematical expressions that represent a specific number or result, such as the calculation of the number of possible combinations using a combinatorial formula. These formulas summarise the solution and often represent the desired result. The *counting process* describes the steps that a student goes through, either mentally or physically, to solve a combinatorial problem. This includes, for example, applying the multiplication principle or systematically breaking a task down into cases. The *sets of outcomes* consist of all possible outcomes or combinations that belong to a specific counting task. For example, the sets of outcomes could represent all possible arrangements of a word. The concept of the solution set helps students visualize what they are counting or calculating and can help them to better understand the structure of a problem. These three components are dynamically related, and the ability to switch between them is crucial for understanding and solving combinatorial tasks (Lockwood, 2013). In one direction, a counting process can lead to the development of a formula/expression. Conversely, an expression can also represent a counting process that a student goes through in their head. Difficulties often arise when students mechanically apply a formula without understanding the underlying counting process. This can lead to errors if they are not aware of the logical structure behind the formula (Lockwood, 2013). Furthermore, a counting process can generate or structure sets of outcomes. Conversely, understanding sets of outcomes can also lead to the selection of a suitable counting process. This relationship helps students to check their solution since organizing the sets of outcomes shows whether all possible results have been considered, and the process was correct. A lack of understanding of this relationship can lead students to overlook or double-count results (Lockwood, 2013). And, ultimately, an expression can directly represent sets of outcomes if a student understands it in terms of a specific structure (e.g., as the number of all possibilities of a group) (Lockwood, 2013).

Following on from this point, in our study 17- to 18-year-old students were asked to formulate a word problem for a given equation for the combinatorial operation arrangement without repetition. Their understanding will be analysed in the context of Lockwood's (2013) model of combinatorial thinking. The analysis of changes between different forms of representation (in this case between symbolic and verbal) can make statements about students' conceptual understanding (e.g., Kilpatrick et al., 2001). The three forms of representation include concrete (everyday) situations (usually described verbally), graphical representations of quantities, and symbolic notations, usually in form of equations corresponding to the underlying quantities and arithmetical operations. A pronounced understanding of operations is demonstrated by the ability to translate back and forth between these different 'languages'. Consequently, six different translation directions are possible.

RESEARCH QUESTION AND METHODOLOGICAL APPROACH

This study aims at expanding the field of research on combinatorics by investigating students' conceptual knowledge of the combinatorial operation arrangement without repetition at the end of their school years and is intended to follow up on a previous study by Thomas and Pöhler (accepted), in which word problems for permutation were analyzed by students of the same age. The following research question will guide this paper:

RQ: What understanding do students show regarding the combinatorial operation arrangement without repetition when formulating their own word problems?

A paper-pencil test was developed to assess the students' conceptual understanding, which consists of two parts: The first part uses a multiple-choice item to ask which of the six combinatorial operations (permutation, arrangement, and combination, both with and without repetition) the students know. Subsequently, the students' conceptual understanding is to be investigated by changing the representation. Therefore, they should formulate a word problem for a given equation. This describes the change from formal representation (given equation) to a concrete (everyday) situation (word problem). The students were given the following task:

Formulate a combinatorial word problem for the following equation:

$$6 \cdot (6 - 1) \cdot (6 - 2) = \frac{6!}{(6 - 3)!} = 120$$

If the students do not answer this question, they can give a reason at the end of the third part of the test why they found it difficult to formulate a word problem. Based on Lockwood (2013), students can understand the following components of the combinatorial thinking model: Students should interpret this formula (variation without repetition) and understand that it calculates the number of ways in which three objects out of six different options can be filled in a particular order. The counting process is not explicitly given, but students must understand it implicitly by developing a suitable word problem. Learners need to understand that the formula is based on a counting process with six options for the first selection; five options for the second

selection (one less as no repetition is allowed) and four options for the third selection. By going through this counting process, students can better understand how the formula is formed and what it describes. The sets of outcomes in this case would be the collection of all possible combinations in which three objects can be selected or distributed from a set of six objects, considering the order. This task of formulating a word problem can help students understand the relationship between formulas/expressions and counting processes, as they must transfer the abstract formula into a real-life context.

One hundred sixty-one students (71 females, 84 males, six no gender) participated in the study. Depending on the type of school, students in Year 12 (Gymnasium) or Year 13 (Gesamtschule) were surveyed. In the federal state of Brandenburg (Germany), there are generally two ways to obtain a university certificate: Either the students attend a Gymnasium and are in school for 12 years, or they attend a Gesamtschule and are in school for 13 years. The students are between 17 and 18 years old. A total of 15 classes were surveyed at six schools, each taught by different teachers. The school classes were selected so that all students had at least four 90-minute blocks of combinatorics lessons in the last school year (according to the teacher responsible for the lessons) and have completed combinatorics lessons according to the curriculum at the time of the test. Whereas at the Gymnasium combinatorics is taught at the end of year 11, it is taught at the end of year 12 at the Gesamtschule. In addition, the teachers were asked about the organization of combinatorics lessons. As the same textbook is used in all classes and, according to the teachers involved, they primarily orientate themselves on this, the students received similar content and lesson time input.

The data was analyzed using a qualitative content analysis according to Mayring (2015). Thereby the categories were determined partly deductively and partly inductively: The categories for word problems for the combinatorial operation permutation (= deductive category building) described in the study by Thomas and Pöhler (accepted) should primarily serve as a basis for the assessment of adequate word problems. Due to the consideration of a different combinatorial operation, these categories were adapted to the arrangement without repetition and, in particular, viable word problems were subdivided into the implicit combinatorial models according to Dubois (1984) (*Selection* problems involve choosing a set of k elements from a set of n elements, with distinctions made based on whether the order of selection matters (permutations or combinations) and whether an element can be selected more than once (with or without repetition) (Batanero et al., 1997b). *Distribution* problems encompass distributing n objects into m cells or assigning n objects to m cells (Batanero et al., 1997a). *Partition* problems involve dividing a set of m objects into n subsets.).

To answer the research question of what understanding students show regarding the combinatorial operation arrangement without repetition, the students' own formulated word problems were categorized deductively and inductively (Mayring, 2015). Based on the 161 word problems formulated by the students, the answers could be

summarized in seven categories – three were formed deductively (C1 – C3) and four inductively (C4 – C7). The **deductive categories** are the following: The participants can, therefore, formulate a word problem for arrangement without repetition (*arrangement word problem C1*), which is based either on the combinatorial model *selection (C1.1)* or *distribution (C1.2)*. Both permutation and arrangement without repetition do not fulfil this requirement for partition, as permutation considers the entire set n without dividing it into subsets, and the focus is on the order of the entire set. This also applies to arrangement - here, too, there is no division of the basic set n into distinguishable classes. In addition, a formulation on the *product rule / multiplication principle (C2)* is possible so that a word problem can be formulated for the equation, which is combinatorial. Furthermore, the participants can write a *word problem for multiplication (C3)*, which is suitable for the equation but is not combinatorial. The following **inductive categories** in Table 1 could be formed on the basis of the data:

Another independent person coded a total of 25% of the data set. The intercoder reliability rate κ was determined using MAXQDA 2024 and was $\kappa=.72$ for the whole category system. It can, therefore, be estimated as substantial (Landis & Koch, 1977). Consensus was subsequently found for the disjunctive cases.

Inductive category	Explanation
C4: Ambiguous combinatorial word problem	C4.1: A word problem is written in which the elements are partially identical.
	C4.2: A word problem is written in which the elements are partially identical and no selection is made.
C5: Probability	The word problem asks for a probability and not the number of elements.
C6: No concrete answer	The field is left blank, a slash is set, or no word problem is written.
C7: Other	The given answer cannot be assigned to the other categories.

Table 1: Inductively categories for the operation arrangement without repetition (N=161)

RESULTS

The 35 adequate word problems are categorized in C1. In contrast to categories C2 and C3 (which are not considered further in this paper), answers are categorized in C1 in which three objects with the same power are arranged or selected from six different objects. The following word problems can be categorized as C1: *Three marbles are selected from an urn containing six different marbles. How many different combinations can be drawn? (C1.1)* or *Six athletes run a race. How many different*

possible arrangements are there on the winner's podium? (C1.2) These two word problems are adequate and show two different combinatorial models. The first word problem is based on a selection, since three marbles are *chosen from* six (a total of 20 students wrote such a word problem), and the second is based on a distribution, since people are *placed on* a podium (a total of 15 students wrote such a word problem). Both answers suggest that the students have understood the counting process underlying the formula. This is because they recognize that the number of possible arrangements decreases by one at each selection stage, which corresponds to the step-by-step approach of the formula. Furthermore, word problems can be categorized that do not fit the arrangement exactly (C4), i.e. word problems in which elements are partially identical (C4.1) and in which, in addition, no selection is made (C4.2). Ten students formulated word problems that cannot be categorized as arrangements with repetition (C4): *Three balls are drawn from a bowl with three white and three black balls. How many different combinations can be drawn?* (C4.1) Although it is clear in this word problem that three marbles are selected from six (selection idea), but several elements are identical, which does not lead to the expected result of 120 possibilities. The following word problem describes a different combinatorial operation: *Determine the number of ways to arrange the letters of the name HANNAH* (C4.2). Similar to the previous word problem, this word problem shows that several elements are identical. Also, no selection of letters is made here, which is necessary for the given equation. In Category C4, both word problems show that students apply the formula mechanically without fully understand what conditions need to be met (e.g. they present the numbers in the term in context, but the relationship is unclear). In the remaining word problems, it is shown that the question is asking about a probability (C5) and not about the number of combinations, or answers are given that provide any content-related answers (C6), e.g. no word problems are written. Finally, there is the category 'other' (C7), which includes word problems that cannot be categorized into the other categories.

DISCUSSION AND CONCLUSION

It is striking that 64% of students (103 out of 161) at the end of school could not formulate a word problem for arrangement without repetition. This is surprising since combinatorics, including factorials, has been part of the German educational standards for mathematics since 2004, typically introduced in the 10th grade (KMK, 2022). Moreover, all teachers interviewed stated that arrangement without repetition is the first combinatorial operation they discuss with students, often using examples like the urn model or a horse race. Students who did not attempt the task cited reasons such as the topic being taught too long ago, general difficulties in writing word problems, or insufficient knowledge of combinatorics. The 35 responses in category C1 reveal valuable insights into students' conceptual understanding of combinatorics, particularly through the lens of Lockwood's (2013) combinatorial thinking model. These students correctly applied the formula to suitable contexts, showcasing an understanding of arranging three elements out of six, where order matters and repetition is excluded. Word problems in C1.1 and C1.2 demonstrate a clear grasp of

sets of outcomes, evident in the explicit specification of order. For example, in C1.2, the sets of outcomes are structured by podium placements, emphasizing the importance of order and variation. This demonstrates flexibility in applying combinatorial principles to various scenarios and highlights an understanding of the mathematical structure behind the tasks. Lockwood's (2013) model helps illustrate the level of conceptual understanding, showing that these students can transfer the formula to contextual applications. Interestingly, textbooks often focus on selection problems (Höveler, 2014), yet similar numbers of students successfully formulated problems for both selection and distribution tasks. In contrast, the 10 responses in category C4 reveal misunderstandings, particularly in interpreting the problem and applying the formula. In C4.1, students ignored the importance of order, treating the task as a combination rather than an arrangement. In C4.2, they addressed order but overlooked the repetition of letters in "HANNAH," leading to an incorrect application of the formula. These issues indicate gaps in understanding the counting process and recognizing the influence of order on the solution. Overall, while students in category C4 placed the formula into a context, they struggled to interpret it correctly and apply it under the appropriate conditions (Lockwood, 2013). Compared to a previous study by Thomas and Pöhler (accepted), which examined permutation without repetition, these results are more concerning. In that study, about half of the 119 students formulated viable word problems, whereas, in this study, only a quarter managed to do so for arrangement tasks. Approximately two-thirds of students left the word problem field blank for arrangement, compared to one-third in the permutation study. Despite these challenges, no responses in this study revealed fundamental mathematical misunderstandings, such as treating the factorial sign as a variable. These findings underscore the need for further research into conceptual understanding, particularly as combinatorics plays a critical role in fostering mathematical thinking and is integral to the curriculum.

Despite the sample size, our study highlights important aspects of students' combinatorial knowledge. With 161 responses, we demonstrate the need to emphasize conceptual understanding, such as formulating word problems, in teaching. While this study focuses on arrangement without repetition, further investigation of other combinatorial operations is essential for broader conclusions. Importantly, the findings underline the need to balance procedural knowledge with conceptual understanding in combinatorics lessons. Encouraging students to create word problems could enhance their grasp of combinatorial operations, fostering deeper comprehension and supporting success in advanced mathematics while promoting rich mathematical thinking.

References

- Batanero, C., Godino, J., & Navarro-Pelayo, V. (1997a). Combinatorial Reasoning and its Assessment. In I. Gal & J. B. Garfield (Eds.), *The Assessment Challenge in Statistics Education* (pp. 239–252). IOS Press.

- Batanero, C., Navarro-Pelayo, V., & Godino, J. D. (1997b). Effect of the Implicit Combinatorial Model on Combinatorial Reasoning in Secondary School Pupils. *Educational Studies in Mathematics*, 32(2), 181–199.
- Dubois, J. G. (1984). Une systématique des configurations combinatoires simples. *Educational Studies in Mathematics*, 1(15), 37–57.
- Fischbein, E., & Gazit, A. (1988). The Combinatorial Solving Capacity in Children and Adolescents. *Zentralblatt für Didaktik der Mathematik*, 5, 193–198.
- Halani, A. (2013). Venn diagrams as visual representations of additive and multiplicative reasoning in counting problems. *Proc. of the 16th Annual Conference on Research in Undergraduate Mathematics Education*, 2, 129–137.
- Hefendehl-Hebeker, L., & Törner, G. (1984). Über Schwierigkeiten bei der Behandlung der Kombinatorik. *Didaktik der Mathematik*, 12(4), 245–262.
- Hiebert, J. & Carpenter, T.P. (1992). Learning and Teaching with Understanding. In D.A. Grouws (Ed.), *Handbook of Research on Mathematics Teaching and Learning: A Project of the National Council of Teachers of Mathematics* (pp. 65–97). Macmillan.
- Höveler, K. (2014). *Das Lösen kombinatorischer Anzahlbestimmungsprobleme: Eine Untersuchung zu den Strukturierungs- und Zählstrategien von Drittklässlern* (Doctoral Dissertation). Universitätsbibliothek Dortmund.
- Kilpatrick, J., Swafford, J., & Findell, B. (2001). *Adding it up: Helping children learn mathematics*. National Academy Press
- KMK – Ständige Konferenz der Kultusminister in der Bundesrepublik Deutschland (Eds.) (2022). *Standards für die Lehrerbildung – Bildungswissenschaften. Beschluss der Kultusministerkonferenz vom 16.12.2004 und vom 04.12.2003, i.d.F. vom 23.06.2022*. KMK.https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2022/2022_06_23-Bista-ESA-MSA-Mathe.pdf
- Lamanna, L., Gea, M. M., & Batanero, C. (2022). Do Secondary School Students' Strategies in Solving Permutation and Combination Problems Change with Instruction?. *Canadian Journal of Science, Mathematics and Technology Education*, 22(3), 602-616.
- Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1), 159.
- Lockwood, E. (2013). A model of students' combinatorial thinking. *The Journal of Mathematical Behavior*, 32(2), 251-265.
- Mayring, P. (2015). Qualitative Content Analysis: Theoretical Background and Procedures. In A. Bikner-Ahsbals, C. Knipping, & N. Presmeg (Eds.), *Approaches to Qualitative Research in Mathematics Education: Examples of Methodology and Methods*, 365–380. Springer
- Thomas, C. & Pöhler, B. (accepted). Analysis of word problems of students at university in relation to the factorial. In *Proc. of the 15th International Congress on Mathematical Education*. ICME.

THE MODERATING ROLE OF SELF-CRITICISM IN THE RELATIONSHIP BETWEEN MATHEMATICS CONFIDENCE AND ANXIETY

Yu-Shan TING¹, Wee Tiong SEAH²

¹National Cheng Kung University, Taiwan

²Faculty of Education, The University of Melbourne, Australia

This study aims to investigate the moderating role of self-criticism in the relationship between mathematics confidence and anxiety among Taiwanese elementary school students. While mathematics confidence negatively correlated with anxiety, self-criticism was positively associated with anxiety. Results also indicate significant gender differences, with boys exhibiting higher confidence and girls reporting greater anxiety. Moderation analysis indicates that self-criticism strengthened the protective effect of confidence on reducing anxiety. These findings highlight the importance of addressing self-criticism and fostering confidence to reduce mathematics anxiety.

BACKGROUND

Students' confidence in mathematics enjoys a strong positive relationship with mathematics achievement at both elementary and secondary school levels in 70 jurisdictions globally (von Davier et al, 2024). According to TIMSS 2019 data, this is no different amongst the East Asian jurisdictions (Wang, King & Leung, 2023). In fact, students' confidence in mathematics in East Asia best predicted their performance (Wang, King & Leung, 2023). Yet, at both Grades 4 and 8, the mathematics confidence levels of these East Asian students were ranked 50th and 27th amongst the 58 and 39 participating jurisdictions respectively (Mullis et al, 2020).

To shed some light on this paradox, the 'Valuing in Confidence' (ViC) study (Third Wave Lab, 2024) was set up in 2024, bringing together research teams from nine jurisdictions across four continents. This paper reports on the progress of the Taiwanese team, which examines a specific aspect of the confidence-achievement relationship, that is, confidence-anxiety. In particular, given that Taiwanese undergraduates reported significantly higher levels of self-judgement, isolation, and overidentification compared to their peers in Thailand and the United States (Neff et al, 2008), self-criticism may undermine students' confidence, intensify anxiety, and hinder learning outcomes in mathematics. As such, the role of self-criticism is examined in this study to help make sense of the paradox mentioned above. Let us first review relevant literature.

CONFIDENCE IN MATHEMATICS

TIMSS 2019 regards students' mathematics confidence as "how well students think they can do mathematics" (Mullis et al, 2020, p. 435). It is similar, but not identical, to

self-efficacy. “Self-efficacy is a belief that is specific to a particular situation, while self-confidence is a general personality trait. Accordingly, self-efficacy can be conceptualized as case-specific self-confidence” (Çiftçi & Yıldız, 2019, p. 684).

Amongst existing approaches to assessing student confidence, the TIMSS 2019 ‘Students Confident in Mathematics’ scale is a 4-point Likert scale asking respondents the extent to which they agree with each of nine statements.

MATHEMATICS ANXIETY

Mathematics anxiety (MA) refers to negative emotional responses related to mathematics participation or thinking, often in the form of worry and nervousness (Richardson & Suinn, 1972), which (can) interfere with doing mathematics (Ashcraft & Kirk, 2001). Thus, MA can impact a student’s mathematics education well-being, and also the learning of mathematics knowledge, skills and competencies. Also well-documented is the inverse relationship between confidence in mathematics and MA. In this sense, in the wider context of the ViC study, the ability to reframe or manage MA (see Buckley & Sullivan, 2023) might mitigate the undesirable effects of low mathematics confidence on mathematics performance.

In other words, given that East Asian students possess lower levels of mathematics confidence compared to their peers elsewhere, might their stellar mathematics performance have been moderated by culturally-unique ways of experiencing, addressing or controlling MA? Given that Taiwan and the other high-performing East Asian jurisdictions share a common Confucian Heritage Culture (CHC) (Biggs, 1996), might there be any Confucian value which interacts with MA? In this study, we were thus motivated to explore the role of the construct of self-criticism.

SELF-CRITICISM

Self-criticism, which emphasizes self-improvement and responsibility, is rooted in Confucianism, as reflected in the Analects 1:4 (Confucius, 2007). It has been found to co-exist with self-compassion amongst Asian Americans (Boyras et al, 2021). Yet, self-criticism can also involve “feelings of unworthiness, inferiority, failure, and guilt” (Warren et al, 2016, p. 19). Thus, self-criticism can serve both adaptive and maladaptive roles. While it can prompt students to reflect on their mistakes and strive for progress, it may also function as an internalized source of stress.

Considering these in the context of the ViC study, the purpose of the Taiwanese study is to investigate the relationship between mathematics confidence and MA, with a specific focus on the moderating role (if any) of self-criticism. By examining how self-criticism influences the protective effects of confidence on anxiety, this study aims to provide insights into the dynamic interplay between cultural and psychological factors in shaping students' emotional experiences in mathematics learning. The following Research Questions guided the conduct of this study:

RQ1: What are the gender differences (if any) in mathematics confidence, mathematics anxiety, and self-criticism among upper elementary school students in Taiwan?

RQ2: What is the relationship between mathematics confidence, mathematics anxiety, and self-criticism?

RQ3: To what extent does self-criticism moderate the effects of mathematics confidence on mathematics anxiety?

RESEARCH FRAMEWORK

The research framework evolves from the literature review above, relating confidence in mathematics and MA, and with self-criticism as a moderating variable. Mathematics confidence is expected to negatively predict MA, reflecting its protective role. However, the strength of this relationship should vary depending on the level of student self-criticism in mathematics learning. High self-criticism may weaken the buffering effect of confidence, while low self-criticism may enhance it.

METHODOLOGY

Participants

The participants were 335 Grades 5 and 6 students randomly selected from 13 classes in two public elementary schools located in northern and southern Taiwan. Two participants did not complete the questionnaire in full. Consequently, the final valid sample consisted of 333 students, made up of 175 boys (52.6%) and 158 girls (47.4%).

Data Collection Instrument

A questionnaire which incorporates the following three inventories was administered to participants. All three inventories used a 4-point Likert scale, with response options ranging from 1 (strongly disagree) to 4 (strongly agree). Demographic variables including student gender were also collected. The questionnaire is accessible at: <https://thirdwavelab.education.unimelb.edu.au/study-12-valuing-in-confidence-vic/>

The Inventory of Mathematics Confidence (IMC), adapted from TIMSS 2019, was employed to assess participants' confidence in mathematics. The IMC consisted of three items selected from the nine used by TIMSS 2019, and demonstrated good internal consistency, with a Cronbach's alpha of .82. An example item is: 'I perform well in mathematics'. Confirmatory factor analysis (CFA) indicated strong factor loadings (.89, .72, and .73). Additionally, the Composite Reliability (CR) was .83, and the Average Variance Extracted (AVE) was .61.

The Inventory of Mathematics Anxiety (IMA), derived from PISA 2022 (OECD, 2023), was designed to assess students' MA. It has six items and demonstrated good internal consistency, with a Cronbach's alpha of .87. An example item is: 'I feel nervous when solving math problems'. CFA indicated an acceptable model fit (CFI = .92, TLI = .86, SRMR = .06). Additionally, the CR was .87, and the AVE was .54.

The Inventory of Self-Criticism in Mathematics (ISCM), developed for this study, was employed to assess participants' self-criticism related to mathematics learning. The ISCM consisted of three items and demonstrated good internal consistency, with a

Cronbach's alpha of .83. An example item is: 'When I perform poorly in mathematics, I criticize myself'. CFA indicated strong factor loadings for the three items (.70, .87, and .79). Additionally, the CR was .84, and the AVE was .61.

Data Collection

The questionnaire was administered by teachers of participating schools in May and June 2024, and student participants filled them in in class. This typically took about 15 minutes. Completed questionnaires were collected, coded, and processed by the first author. All responses were anonymized to protect participants' privacy.

Data Analysis

The data were analysed using SPSS 21 and AMOS 26. SPSS was used to assess reliability through Cronbach's alpha, perform descriptive statistics, ANOVA, and conduct moderation analysis using PROCESS Model 1 (Hayes, 2018). AMOS was employed to conduct CFA and to calculate CR and AVE to ensure construct validity.

RESULTS

Preliminary Analysis

The results showed that the Taiwanese Grades 5 and 6 students reported moderately low levels of self-criticism in mathematics learning ($M = 2.14$, $SD = 0.91$), of mathematics confidence ($M = 2.37$, $SD = 0.88$), and of MA ($M = 2.24$, $SD = 0.82$), since the mean values were all below the midpoint of 2.5 on the 4-point scale.

A one-way ANOVA was conducted to examine gender differences in mathematics confidence, MA, and self-criticism. The results (Figure 1) revealed significant gender differences in both mathematics confidence and MA. For mathematics confidence,

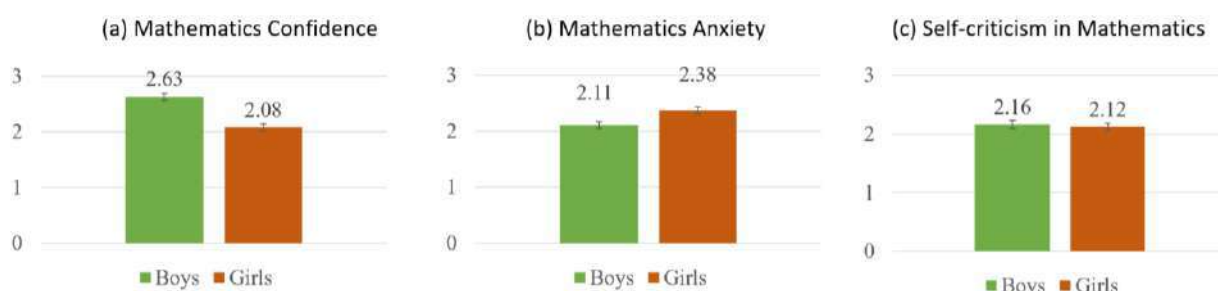


Figure 1: Gender differences.

$F(1, 331) = 35.95$, $p < .001$, $\eta^2_p = .10$. Boys reported significantly higher mathematics confidence ($M = 2.63$, $SE = .06$) compared to girls ($M = 2.08$, $SE = .07$). The results also showed a significant gender differences on MA, $F(1, 330) = 9.01$, $p = .003$, $\eta^2_p = .03$. Girls reported significantly higher MA ($M = 2.38$, $SE = .77$) compared to boys ($M = 2.11$, $SE = .84$). Finally, there were no significant gender differences in self-criticism in mathematics learning, $F(1, 331) = .14$, $p = .71$, $\eta^2_p = .00$.

Correlation Analysis

A Pearson correlation analysis was conducted to examine the relationships among mathematics confidence, MA, and self-criticism in mathematics learning. The results indicated a significant negative correlation between mathematics confidence and MA ($r = -.44, p < .001$), suggesting that higher mathematics confidence is associated with lower MA. Additionally, MA was positively correlated with self-criticism in mathematics ($r = .36, p < .001$), indicating that higher levels of MA are associated with higher levels of self-criticism. However, there was no significant correlation between mathematics confidence and self-criticism in mathematics learning ($r = -.04, p = .46$).

Moderation Analysis

A moderation analysis was conducted using Hayes' PROCESS Model 1 (Hayes, 2018) with bias-corrected bootstrap 95% confidence intervals based on 5000 bootstrap samples. MA was the dependent variable, mathematics confidence the independent variable, and self-criticism was included as the moderator. Gender was added as a covariate to ensure that the moderation effect was not confounded by gender differences.

Regarding the main effects, mathematics confidence was not a significant predictor of MA ($b = -.17, SE = .11, t = -1.59, p = .11$). However, self-criticism was positively associated with MA ($b = .54, SE = .11, t = 4.84, p < .001$). Regarding the interaction effect, the interaction between mathematics confidence and self-criticism ($b = -.10, SE = .04, t = -2.27, p = .024$) was significant, indicating that self-criticism moderated the relationship between mathematics confidence and MA.

At all levels of self-criticism, mathematics confidence significantly predicted lower MA ($p < .001$) (see Table 1). Notably, the effect size increased with higher levels of self-criticism, suggesting that self-criticism amplifies the protective effect of confidence on anxiety.

Level	b (SE)	t	p
Low (1)	-0.27 (0.07)	-3.90	<.001
Moderate (2)	-0.37 (0.05)	-8.01	<.001
High (3.24)	-0.49 (0.06)	-7.57	<.001

Table 1: Conditional effects of mathematics confidence on mathematics anxiety at different levels of self-criticism.

DISCUSSION

This study examined mathematics confidence, MA, and self-criticism in mathematics learning among Taiwanese students. Results showed that students reported low levels of mathematics confidence, reflecting TIMSS 2023 (von Davier et al, 2024) in which Taiwanese students at both elementary and junior high levels showed lower mathematics confidence compared to the international average (von Davier et al, 2024). On the other hand, the students also reported low levels of MA and self-criticism, when Taiwanese students have been known internationally to be relatively

more anxious and self-critical in mathematics learning (Heine, 2003; Neff et al, 2008; OECD, 2023). Might a societal valuing of *humility* explain the low confidence? Might the under-reporting of MA and self-criticism be due to cultural valuing of, say, *perseverance* and *diligence*?

In relation to RQ1, our data revealed significant gender differences in mathematics confidence and anxiety, with boys and girls reporting higher confidence and higher anxiety respectively. These findings align with prior research (e.g. Hannula et al, 2005), suggesting that little has improved over the last decades. Societal, parents and teachers' gender-biased expectations might still be contributing to girls' higher anxiety and lower confidence in mathematics (Gunderson et al., 2012). Interestingly, the lack of a statistically significant difference in self-criticism between boys and girls may be attributed to its pervasiveness as a Confucian value in Taiwanese society.

As for RQ2, correlation analysis demonstrated a significant negative relationship between mathematics confidence and anxiety, and a positive relationship between anxiety and self-criticism. The former aligns with prior research showing that higher levels of MA are associated with lower mathematics confidence (Ashcraft & Moore, 2009). The latter reflects a pattern where negative self-appraisals heighten emotions (Beck, 1976). The lack of a significant correlation between confidence and self-criticism implies that self-criticism may be influenced more by cultural expectations than by perceived competence (Neff et al., 2008). Its significance and role in CHCs such as Taiwan are thus worthy of further examination and discussion.

Turning to RQ3, the moderation analysis revealed that self-criticism significantly moderated the relationship between mathematics confidence and anxiety, with the negative relationship being stronger at higher levels of self-criticism. In other words, for Taiwanese students who were more self-critical, their mathematics confidence would help to reduce anxiety more effectively. This may be explained by the cultural context of East Asian societies, in which self-criticism is often viewed as a form of self-improvement rather than purely negative self-evaluation (Heine, 2003). In contrast to Western cultures, which emphasize positive self-regard, CHCs prioritize self-reflection, encouraging individuals to recognize their shortcomings and strive for improvement. This demonstrates an adaptive function of self-criticism as often experienced by East Asians, when it also acts as a moderator that amplifies the protective effect of confidence. At the same time, while this cultural emphasis on self-criticism may foster perseverance and diligence, it may also have unintended emotional consequences, particularly when students are unable to meet high personal standards. Indeed, self-criticism can increase the risk of anxiety. As such, we have found that for the Taiwanese students, self-criticism is both a risk factor and a moderator that amplifies the protective effect of confidence. Since the current study has not been designed to examine how these variables work together, future research may be designed to deepen our understanding of findings such as the negative MA – mathematics confidence relationship being stronger at higher levels of self-criticism.

These findings certainly have implications for mathematics classroom practice. Given that self-criticism amplifies MA, schools can work with parents in supporting students with high self-criticism. Such support can involve the teaching of self-compassion and of positive self-talk. Also, since the protective effect of students' mathematics confidence becomes more pronounced at higher levels of self-criticism, schools can consider introducing programs aimed at developing mathematics confidence amongst their students. Personalised encouragements, for example, can play a role here.

From the perspective of the wider ViC study, the findings of this study contribute to our understanding of how the relatively more self-critical Taiwanese – or East Asian more generally – students with relatively low confidence have been performing better in mathematics assessments.

REFERENCES

- Ashcraft, M. H., & Kirk, E. P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology: General*, 130(2), 224–237. <https://doi.org/10.1037/0096-3445.130.2.224>
- Ashcraft, M. H., & Moore, A. M. (2009). Mathematics anxiety and the affective drop in performance. *Journal of Psychoeducational assessment*, 27(3), 197–205. <https://doi.org/10.1177/0734282908330580>
- Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. International Universities Press.
- Biggs, J. (1996). Western misconceptions of the Confucian-heritage learning culture. In D. A. Watkins & J. B. Biggs (Eds.), *The Chinese learner: Cultural, psychological and contextual influences* (pp. 45–68). ACER.
- Boyratz, G., Legros, D. N., & Berger, W. B. (2021). Self-criticism, self-compassion, and perceived health: Moderating effect of ethnicity. *Journal of General Psychology*, 148(2), 149–167. <https://doi.org/10.1080/00221309.2020.1746232>
- Buckley, S., & Sullivan, P. (2023). Reframing anxiety and uncertainty in the mathematics classroom. *Mathematics Education Research Journal*, 35(1), 157–170. <https://doi.org/10.1007/s13394-021-00393-8>
- Çiftçi, S. K., & Yildiz, P. (2019). The effect of self-confidence on mathematics achievement: The metaanalysis of Trends in International Mathematics and Science Study (TIMSS). *International Journal of Instruction*, 12(2), 683–694. <https://doi.org/10.29333/iji.2019.12243a>
- Confucius (2007). *The Analects of Confucius* (B. Watson, Trans.). Columbia University Press (Original work published 220).
- Gunderson, E. A., Ramirez, G., Levine, S. C., & Beilock, S. L. (2012). The role of parents and teachers in the development of gender-related math attitudes. *Sex roles*, 66, 153–166. <https://doi.org/10.1007/s11199-011-9996-2>

- Hannula, M. S., Maijala, H., Pehkonen, E., & Nurmi, A. (2005). Gender comparisons of pupils' self-confidence in mathematics learning. *Nordic Studies in Mathematics Education*, 3–4, 29–42. <https://doi.org/10.7146/nomad.v10i3-4.147166>
- Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2nd ed.). New York: Guilford Press.
- Heine, S. J. (2003). An exploration of cultural variation in self-enhancing and self-improving motivations. In V. Murphy-Berman & J. J. Berman (Eds.), *Cross-cultural differences in perspectives on the self* (pp. 118–145). University of Nebraska Press.
- Mullis, I. V. S., Martin, M. O., Foy, P., Kelly, D. L., & Fishbein, B. (2020). *TIMSS 2019 international results in mathematics and science*. IEA.
- Neff, K. D., Pisitsungkagarn, K., & Hsieh, Y.-P. (2008). Self-compassion and self-construal in the United States, Thailand, and Taiwan. *Journal of Cross-Cultural Psychology*, 39(3), 267–285. <https://doi.org/10.1177/0022022108314544>
- OECD (2023) *PISA 2022 results (Volume I): The state of learning and equity in education*. OECD Publishing. <https://doi.org/10.1787/53f23881-en>
- Richardson, F. C., & Suinn, R. M. (1972). The mathematics anxiety rating scale. *Journal of Counseling Psychology*, 19, 551–554.
- Third Wave Lab (2024). *Study 12: Valuing in Confidence [ViC]*. <https://thirdwavelab.education.unimelb.edu.au/study-12-valuing-in-confidence-vic/>
- von Davier, M., Kennedy, A., Reynolds, K., Fishbein, B., Khorramdel, L., Aldrich, C., Bookbinder, A., Bezirhan, U., & Yin, L. (2024). *TIMSS 2023 international results in mathematics and science*. Boston College, TIMSS & PIRLS International Study Center. <https://doi.org/10.6017/lse.tpisc.timss.rs6460>
- Wang, F., King, R. B., & Leung, S. O. (2023). Why do East Asian students do so well in mathematics? A machine learning study. *International Journal of Science and Mathematics Education*, 21(3), 691–711. <https://doi.org/10.1007/s10763-022-10262-w>
- Warren, R., Smeets, E., & Neff, K. (2016). Self-criticism and self-compassion: Risk and resilience: being compassionate to oneself is associated with emotional resilience and psychological well-being. *Current Psychiatry*, 15(12), 18–28.

ESTABLISHING PRACTICES OF UNPACKING MEANINGS OF MATHEMATICAL CONCEPTS IN MULTIPLE LANGUAGES

Ángela Uribe¹ & Susanne Prediger^{2,3}

¹St.Gallen University of Teacher Education, Switzerland, ² TU Dortmund University, Germany, ³IPN Leibniz Institute for Science and Mathematics Education, Germany

Unpacking and comparing the literal meanings of phrases for mathematical concepts in multiple languages can unfold productive discussions to deepen students' conceptual understanding. However, as students often translate only by overall sense, productive practices of language comparison must first be established. The design research study explores a multilanguage-responsive approach with prepared phrases and literal meanings from multiple languages, for the case of delineating area and perimeter. Qualitative analysis of two episodes of whole-class discussions revealed four practices of language comparison with varying degrees of contribution to meaning-making. The analysis indicates that the prepared phrases can support to establish productive practices that students can later transfer to their own home languages.

INTRODUCTION: PRACTICES OF LANGUAGE COMPARISON AND THEIR (PARTIALLY UNEXPLOITED) EPISTEMIC POTENTIAL

Comparative linguistic analyses revealed that different languages can express mathematical concepts with slightly different conceptualizations (Barton, 2008). For example, the English expression “three fifths” for fractions emphasizes the counting of new units “fifths”. Meanwhile, the Turkish expression “beşte üç” begins with the whole “beş” (five), followed by the suffix “te” (therein) and the part “üç,” carrying the whole-part structure (“five-therein three”) more explicitly than English (Prediger et al., 2019).

In mathematics classrooms, the practice of comparing phrases from multiple languages for the same mathematical concept can therefore be productive to deepen students' conceptual understanding. This can be achieved when students realize and connect different concept elements for the concept in view, which together form its holistic meaning (Prediger et al., 2019). While this epistemic potential has been exemplified in some case studies (Barwell, 2018; Prediger et al., 2019), further classroom observations revealed that the epistemic potential is often left unexploited, resulting in missed learning opportunities (Prediger & Uribe, 2021; Ferrari et al., 2023). To address this, we developed a multilanguage-responsive instructional approach with prepared phrases from multiple languages and their literal translations. This approach aims to establish language comparison practices that can later be spontaneously extended to other languages (Prediger & Uribe 2021; Schüler-Meyer et al., 2023). The paper presents a sub-study within a design research project that aims to substantiate this approach for the

case of delineating the measurement concepts of area and perimeter. We qualitatively analyze classroom discussions with respect to the following research question:

What practices of language-induced reflection do students engage in when interpreting phrases for area and perimeter from multiple languages, including their home languages, and how do these practices contribute to meaning making?

THEORETICAL BACKGROUND

Delineating area and perimeter as a typical conceptual challenge

Perimeter and area are measurement concepts for two-dimensional figures. Perimeter refers to the *length around* the outer boundaries of a figure (e.g., measured in standard units of 1 cm), and the area refers to the *space inside* (measured by covering or tiling it with two-dimensional standard units, e.g., squares of 1 cm²). For students, “one of the most persistent findings related to learning to measure area is students’ struggle to distinguish area from [perimeter]” (Smith & Barrett, 2017, p. 364).

In the empirical part of this paper, we explore to what extent unpacking the phrases used to express the concepts of area and perimeter can enhance students’ ability to sustain a clear delineation of the meanings of two distinct concepts: For example, in German language, “Umfang” (perimeter) includes the prefix “um” meaning “around” in “around-catch”, while “Flächeninhalt” (area) the prefix “in”, referring to “inside”, as “area-content-inside”). Figure 1 provides additional examples from other languages.

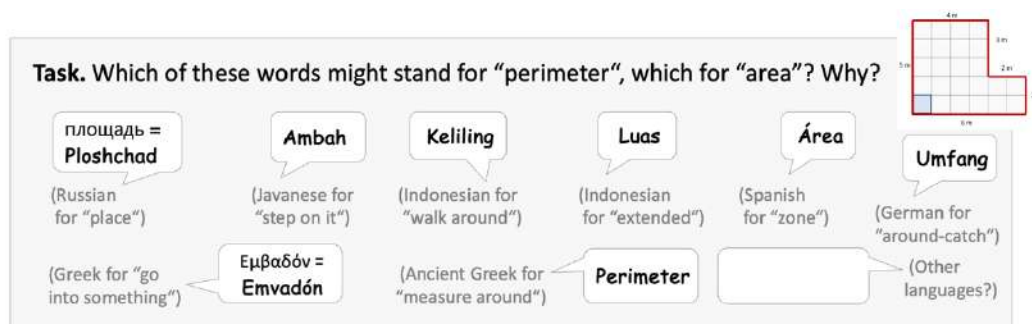


Figure 1. Language reflection for deepening understanding of area and perimeter

Documented challenges in exploiting epistemic potentials of language reflection

From some classrooms, compelling cases were presented how multilingual students drew upon their home languages to unpack the literal meanings of other mathematical concepts and compare the literal translations to connect and delineate different meanings (Barwell, 2018; Prediger et al., 2019). However, in many other classrooms, students’ language awareness was less developed, for instance a Greek-speaking student might translate “Emvadón” simply as “area”, conducting translations by overall sense rather than a word-for-word literal translation. Teachers who do not speak the language in view cannot unpack the more subtle nuances of the literal translation of “go into something” (Prediger & Uribe, 2021). So, the productivity of language comparison

practices seems to depend on a) the phrase chosen from the foreign language, b) the kind of references made for interpreting the foreign phrase into German, c) the explanation of what the German translation means, and d) the addressed concept elements, which might enrich the meaning of the underlying concept.

Design of prepared phrases with literal translations from multiple languages

In order to overcome these well-documented challenges, we have developed tasks to compare and delineate *prepared* phrases from multiple languages together with their literal translations (Schüler-Meyer et al., 2023). These tasks were intended to support teachers to *establish productive language comparison practices* and thereby provide learning opportunities for students to experience what the epistemic potential could be.

This design approach is theoretically based on a definition of (mathematical or discursive) *practice* as recurrent ways of acting mathematically or discursively, which are co-constructed in the classroom interaction (Cobb et al., 2001), and thereby might create certain conceptual learning opportunities, in our case for meaning making on area and perimeter. While tasks can pose certain demands, new practices only emerge in the interaction, which is nurtured by students' contributions and supported by teachers' prompts and scaffolding. Within this theoretical perspective on learning, students' participation in collectively enacted (and supported) practices can lead to more independent accomplishments. Yet, empirical investigations must be conducted to concisely characterize productive and unproductive practices, their factual contributions to meaning-making and teachers' means to establish and support students' engagement in these productive practices.

METHODS

Methods of data gathering in design experiments. The research question that guides our current study is pursued within the research context of the larger design research project ML², which develops and investigates learning environments for exploiting multilingual resources in Grade 5–8 mathematics classrooms (Schüler-Meyer et al., 2023). Typical for the design research methodology is the dual aim of (a) developing and optimizing learning opportunities (in this paper, for meaning making through multiple language comparison), and (b) generating deep insights into the initiated learning processes (in this paper, the language comparison practices and their potentials for meaning making) (Cobb et al., 2001). Design experiments were conducted in three Grade 5–7 classes (with 10-13-year-old students) by research-near teachers new to the classes. The experiments lasted 90 minutes each, of which 30-40 minutes treated the concepts of area and perimeter in view of this paper. In the empirical part, we present the analysis of two whole-class episodes on the task in Figure 1 and the transfer to students' own home languages. The videorecorded episodes were fully transcribed in German.

Methods of qualitative data analysis. The qualitative analysis for the two episodes was conducted using the original German transcript. Afterward, the transcript was translated into English, taking care of preserving the language subtleties in view from the original transcript. For each student contribution, we coded with the analytic scheme in Figure 2, a) which phrase in a foreign language students reflected on, b) to what they referred in its interpretation, c) into what phrase in German they translated it and how they explained it (if they did), and from this, we derived interpretatively d) what concept elements they addressed through the language reflection, which might contribute to enriching the meaning (see Figure 3 for examples). By comparing different moments in the data, we identified different practices of language reflection and their potential for meaning making.

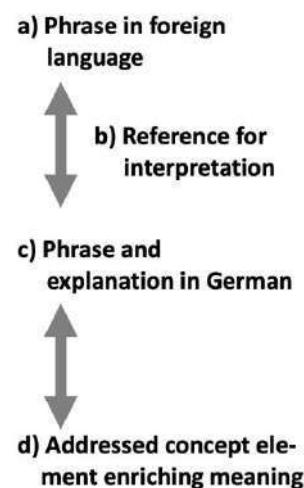


Figure 2.
Analytic scheme

EMPIRICAL INSIGHTS: TWO EPISODES OF LANGUAGE REFLECTION

Episode 1: From associating world knowledge to unpacking given literal translations

The first episode stems from a Grade 5 class. The students discuss which phrases refer to perimeter and which to area, based on the task from Figure 1.

- 10 Paul: I think that [*the Spanish phrase*] “Área” comes [*translates to “Fläche”*] because of the [*US*] word “Area 51”.
- 11 Teacher: Ok, yes, you would have derived it from English.
- 12 Sina: Yes, it [*the literal translation*] says “zone”, and the area is such a zone?
- 13 Teacher: Yes, exactly, great, great explanation.
- 14 Lisa: I have one more word. In Indonesian [*“keliling”*] I have “walk around” at the bottom, that's why I thought of perimeter? Because it's also like in the figure, like this [*gestures around the figure*].

Paul (in T. 10) refers to the Spanish phrase “Área”, associating it with the US term “Area 51” by drawing on world knowledge and cultural references related to the secret military space. The teacher (in T. 11) explicitly acknowledges that this derivation comes from English, reinforcing the linguistic connection. While this world knowledge association is productive for exploring cross-linguistic similarities, it does not necessarily contribute to a deeper understanding of the mathematical concept of area itself.

Sina (in T.12) builds on this discussion by referencing the literal translation “zona” (zone), suggesting that area can be understood as a type of zone, so her contribution becomes more mathematically relevant. While the teacher reinforces her practice of referring to the literal translations given below the speech bubble in the worksheet (T.13), the conceptual implications of “zone” remain unexplored.

Lisa refers to the Indonesian phrase “keliling” for perimeter and immediately connects it to its literal translation (T.14). *Lisa*’s focus on literal translation may have been encouraged by the teacher’s earlier reinforcement of *Sina*’s contribution (T. 12/13). In this way, language reflection by unpacking literal meanings is introduced and practiced. However, it is *Lisa*’s second sentence by which this moment becomes particularly productive for her meaning making: She seems to realize that the term closely aligns with her mental representation of perimeter, which she visually explains by gesturing around the figure. It is this connection from the literal translation “walk around” to this gesture that bears most epistemic potential.

Episode 2: From unpacking given literal translations to own home languages

The second episode stems from a Grade 7 whole-class discussion summarizing what students have individually examined in the given provided phrases for area and perimeter (Figure 1). The teacher prompts students to identify phrases they consider inaccurate for these concepts.

- 31 Cornelius: [*Refers to “Ambah”, Javanese for “step on it”*] So, I find it’s [*referring to “on”*] ... confusing. Perhaps one should try to understand it more as entering **into** an area or simply being inside.
- 32 Teacher: ... In #Greek [*“Emvadón”, Greek for “go into something”*] that is included. Would it be enough then to simply enter the area?
- 35 Rafael: No, because one still needs to move within the area, so that's just too small.
- 37 Rafael: Walk around
- 38 Teacher: Walk around. But walk around was perimeter, right?
- 39 Rafael: Yes, but you can walk around within a surface. [...] Inside
- 44 Teacher: How does that look in Ukrainian? [*Teacher asks Svenja*].
- 45 Svenja: In Ukranian, area is called “площа” (ploscha), which ... means “place”.
- 48 Teacher: What does that have to do with determining area? [*3 sec. break*] Place, area?
- 52 Denise: Mm, so, area in Bosnian is “površina”, and it means “area”.
- 58 Lena: [*referring to question in T. 48*] The area occupies a place.

Cornelius (in T. 31) articulates his concerns that the phrase “Ambah” (Javanese for “area” meaning “step on it”) is confusing, not in delineation to perimeter, but in distinguishing ideas of “on” and “in”. Instead, *Cornelius* suggests a more fitting interpretation, such as “entering into” or “being inside”. The teacher (T. 32) highlights that this idea of “going into something” aligns with the Greek phrase and asks whether it sufficiently conveys the meaning of area. *Rafael* (in T. 35-39) disagrees, referring to a more dynamic idea of moving within the area. However, he does not explicitly express the idea of moving to cover the entire area, leaving it implicit. The teacher (T. 38) uses *Rafael*’s phrase, “walk around”, to initiate a conceptual contrast, asking whether this description might better fit the perimeter rather than the area. *Cornelius* and *Rafael*

refer to the given phrases in different languages and their literal translations. In both cases, this contributes to their meaning making as they distinguish between “on” and “in”, i.e., covering with standard units versus tiling with standard units.

To further mobilize the multilingual resources of the students, the teacher asks *Svenja* for an ad hoc translation into Ukrainian. Although such spontaneous practices of semantic unpacking do not always succeed, *Svenja*, who recently arrived from Ukraine, is able to provide one, as she has attended math classes in Ukrainian in her home country. She translates the literal meaning as “place”, possibly inspired by the similar translation provided in Russian. In order to exploit this phrase for students’ meaning making, the teacher engages also other students in discussions about meanings of the literal translation by asking: “What does that have to do with determining area? Place, area?” (T. 48). At this point, an interesting discussion over 13 turns occurs (non-printed due to space restrictions), which finally ends in the highly condensed and deep explanation by *Lena* (a student from a monolingual family), summarizing: “The area occupies a place”. (T. 58), in which “occupy” offers a third element of meaning for determining areas, “stepping on” and “going inside”.

During the discussion, *Denise* (T. 52) introduces her home language, Bosnian, yet only by translating by overall sense in which the Bosnian word is assigned to the German label without a one-to-one word equivalent or further unpacking. Neither *Denise* nor the teacher were able to quickly structurally unpack the word: According to our inquiry with Bosnian-speaking adults, *po-vrš-ina* seems to be composed of the prefix “po” for “on” or “above”, the component “vrš” stands for “upper layer” and the suffix “ina” is often used for constructing nouns describing properties or physical expanse, together “the upper expanse”. This chance for deeper conceptual exploration was missed.

DISCUSSION

Outcome 1: Replication that language comparison can be applied to initiate meaning-making processes, also for delineating area and perimeter. The multilanguage-responsive approach to unpack and compare phrases from multiple languages has been shown to bear epistemic potential for enhancing students’ understanding of fractions (Bartolini-Bussi et al., 2014), proportional reasoning (Prediger & Uribe, 2021) and algebra (Schüler-Meyer et al., 2023; Ferrari et al., 2023). Our qualitative analysis revealed insights how the potential can also be exploited for delineating area and perimeter, an often-documented student challenge (Smith & Barrett, 2007), with multilingual and monolingual students who do not share all their home languages.

Outcome 2: Disentangle four language comparison practices with increasing potential to deepen understanding. Beyond these replications, the main contribution of the current study is to disentangle how productive practices can be characterized, through the components of the analytic scheme (Figure 2). In the analysis of two short episodes, we identified four practices (Figure 3): *translating by overall sense* was

shown not to contribute to deepening conceptual understanding, whereas *associating world knowledge* (such as “Area 51”) could have contributed if the teacher had leveraged this idea by unfolding the “secret inside” into a good mnemonic bridge. The *intended practice of semantic unpacking* was found in four moments, each of which made additional contributions to deepening conceptual understanding. Three of these moments emerged with prepared phrases (Sina T. 12, Lisa T. 14, Cornelius & Rafael T. 31-35), indicating that the design approach seems to bear the intended effects. The fourth moment (Svenja in T. 45-58) provides first indications that students' transfer from prepared phrases in given languages to their own language is possible, but becomes richer for deepening conceptual understanding when the teacher supports the explicit connection to the mathematical meaning.

Practices of language reflection for meaning making	Translating by overall sense	Associating world knowledge	Semantic unpacking	Structural unpacking
Phrase in foreign language	Denise (T. 52) (Own home language) Bosnian “Površina”	Paul (T. 10) (Given foreign language) Spanish „Area“	Cornelius (T. 31-35) (Given foreign language) Javanese “Ambah”	Missed chance for Denise (T. 52) Bosnian “Površina”
Reference for interpretation	None (only assigning labels)	World knowledge “Area 51”	Literal translations “step on it”	Grammatical structure po-vrš-ina prefix suffix for “the upper expanse”
Phrase and explanation in German	Flächen“inhalt” (Area) Only technical phrase without unfolding	“Fläche” (a particular area)	Not only “on” but “entering into an area” or “being inside”	“Obere Ausdehnung” “The upper expanse”
Addressed concept element enriching meaning	No new facet	Highly debated space for which the inside is secret	Area is not sur-face, but inside	Area as sur-face in its whole extent

Figure 3. Analytic outcome: Four language comparison practices with increasing potential to meaning-making (3 more moments of semantic unpacking unprinted)

The *fourth practice of structurally unpacking phrases*, however, was still on the list of missed chances. While a structural unpacking of the self-introduced Bosnian construction with prefix and suffix would have been highly interesting to reflect on, Denise only offered a translation by overall sense. In other transcripts, we found the first seeds of structural unpacking, but also the need to provide more scaffolding for it.

Limitations. Two short episodes provide only a small empirical base, our findings are further substantiated by analyses of additional classroom videos from the larger project, where we found these four practices as repeated pattern. As the contribution to meaning making can only be inferred from students' explicit contributions, we miss insights into silent meaning making thoughts of students who do not participate verbally. To address this limitation, future studies should combine the analysis of whole-class discussions with more individualized data gathering, e.g., in pair work.

Conclusion and design implications. In summary, we disentangled the differences between *semantic unpacking* (semantically drawing on the components of the words) and *structural unpacking* (breaking down the phrases into prefixes and suffixes and analyzing their structural, word-forming grammatical functions). While semantic unpacking was well initiated, structural unpacking rarely occurred. Therefore, we suggest that in a next project, we should also support teachers in the process of structural unpacking. While the current material has successfully encouraged teachers and students to discuss semantic aspects of different literal translations, future designs should also sensitize students to the possibility of structural unpacking the grammar in prefixes and suffixes, conjunctions, prepositions, as these are the most interesting connectors through which different mathematical relations are expressed.

Funding. The project ML² – Exploiting the Power of Multiple Languages for Mathematics Learning is funded by the European Union (grant no. 2021-1-NL01-KA220-SCH-000024585). We thank our project partners Alexander Schüler-Meyer, Shintia Revina, Tamsin Meaney, Silke Lekauss and Yasemin Platen for teaching the approach.

References

- Bartolini Bussi, M.G., Baccaglini-Frank, A., & Ramploud A. (2014). Intercultural dialogue and the history and geography of thought. *For the Learning of Mathematics*, 34(1), 31–33.
- Barton, B. (2008). *The language of mathematics: Telling mathematical tales*. Springer
- Barwell, R. (2018). From language as a resource to sources of meaning in multilingual mathematics classrooms. *Journal of Mathematical Behavior*, 50, 155–168.
- Cobb, P., Stephan, M., McClain, K., & Gravemeijer, K. (2001). Participating in classroom mathematical practices. *The Journal of the Learning Science*, 10, 113–163.
- Ferrari, E., Meaney, T., & Lekauss, S. (2023). The complexity of task design for utilising the epistemic potential of multiple languages in developing pattern understandings. In P. Drijvers et al. (Eds.), *Proceedings of CERME 13* (pp. 1585-1592). ERME. /hal-04393040.
- Prediger, S., & Uribe, Á. (2021). Exploiting the epistemic role of multilingual resources in superdiverse mathematics classrooms. In A. Fritz, E. Gürsoy, & M. Herzog (Eds.), *Diversity Dimensions in Mathematics and Language Learning* (pp. 80–97). De Gruyter.
- Prediger, S., Kuzu, T., Schüler-Meyer, A., & Wagner, J. (2019). One mind, two languages—separate conceptualisations? *Research in Mathematics Education*, 21(2), 188–207.
- Schüler-Meyer, A., Meaney, T., Uribe, Á., & Prediger, S. (2023). Design heuristic for generating conceptual learning opportunities through multiple languages. In A. Schüler-Meyer, J. Ingram & K. Erath (Eds.), *Proceedings of ETC 12 on Language in the Mathematics Classroom* (pp. 102–110). ERME. hal.science/hal-03992500
- Smith, J. P., & Barrett, J. E. (2017). Learning and teaching measurement: Coordinating quantity and number. In J. Cai (Ed.), *Compendium for Research in Mathematics Education* (pp. 355–385). National Council of Teachers of Mathematics.

EXTRAMATHEMATICAL CONNECTIONS RECOGNIZED BY PROSPECTIVE ELEMENTARY SCHOOL TEACHERS

Juan Pablo Vargas Herrera¹, Yuly Vanegas²

¹Universidad de las Américas; ²Universitat de Barcelona

Research on connections in the context of teacher education is a growing topic of interest in Mathematics Education. However, such research has focused on intra-mathematics type connections and in Secondary Education. This report describes one type of extra-mathematical connection that emerges in the development of a professional task, which deals with geometric notions, posed to 250 future elementary school teachers. Tools of the onto-semiotic approach are used in the analysis. The results show the emergence of the materialization connection, through the analysis of the solution trajectories and the semiotic functions established by the future teachers when they try to design geometric school tasks. Finally, elements for the promotion of better connections and their implications in teacher education programs are discussed.

INTRODUCTION

According to the NCTM (2000) mathematical learning is deeper and more durable if students can connect mathematical ideas with the world around them. From this perspective, numerous authors have investigated mathematical connections (Businkas, 2008; García-García and Dolores-Flores, 2018; Rodríguez-Nieto et al., 2021, among others) establishing definitions, interpretations and categories of analysis for different educational levels. In this research, connections are understood as “a cognitive process through which a person relates two or more ideas, concepts, definitions, theorems, procedures, representations and meanings to each other, to other disciplines or to real life” (García-García and Dolores-Flores, 2018, p. 22).

The study on connections, in the framework of Primary Education teacher training, is a recent field and, therefore, with quite a few unexplored aspects. According to García-García & Dolores-Flores (2018) research on connections has focused mainly on secondary mathematics teachers and focuses on the study of intra-mathematics connections. Therefore, it is necessary to develop research that explores connections in the field of Primary Education teacher training, in order to promote key aspects for their establishment in teaching from the first educational levels.

The main objective of this research report is to evidence and describe a type of extra-mathematical connection that a group of 250 future teachers of Primary Education establish when developing professional tasks related to geometry. The emergence of this connection is justified using the tools of the Onto-semiotic approach. The reported result is part of a doctoral research that detected six different types of extra-

mathematical connections; however, for reasons of length, this report presents the detailed analysis of one of these connections (materialization). It is hoped that the presentation of this evidence will serve as a basis for further theorizing about extra-mathematical connections and, according to García-García (2019), encourage “building and validating a framework for studying mathematical understanding from mathematical connections” (p. 132).

THEORETICAL FRAMEWORK

In the following, some aspects of the mathematical connections are presented, followed by a synthesis of some tools of the Onto-semiotic Approach used in the analyses.

MATHEMATICAL CONNECTIONS

Mathematical connections have been the subject of study in mathematics education, as they underline the importance of linking mathematics with the real world, leading educational research during the last decades (Businkas, 2008). Different agents have urged their use and inclusion in national curricula and in general they have been constituted as key elements for the teaching of mathematics. Despite extensive research, there is no universally accepted definition or categorization of these (García-García, 2019). Several investigations such as those of Businkas (2008) and Gamboa and Figueiras (2014) have defined and typified mathematical connections. In such works they refer to the relationships that can be established between elements inside and outside mathematics. An intra-mathematical connection is identified as one that occurs between elements of mathematics; while a connection that establishes relationships between an object external and one internal to mathematics is defined as extra-mathematical. Regarding extra-mathematical connections, Vanegas and Giménez (2018) propose a categorization when studying the productions of a group of future elementary education teachers. The connections identified were modelling, mediating, semiotic, metaphorical, generic interdisciplinary and materialization. This last type of connection refers to the relationships that allow an extra-mathematical element to materialize a mathematical idea.

TOOLS OF THE ONTO-SEMIOTIC APPROACH (OSA): EPISTEMIC CONFIGURATIONS

The OSA considers that a mathematical practice is any sequence of actions, subject to mathematical rules, performed by someone to solve mathematical problems, communicate its solution to others, validate it, generalize it and use it in other contexts (Godino et al., 2007). It normally becomes relevant not as an individual element, but as part of a system of practices that manifests itself when a person is confronted with a problem situation.

Godino et al. (2007) indicate that to carry out a mathematical practice and subsequently interpret its results as satisfactory, it is necessary to put into operation a series of knowledge. Thus, whenever a mathematical practice is carried out to solve a problem, a series of objects are mobilized, such as situations, problems, concepts, propositions

and procedures, which are involved in the elaboration of arguments that can determine the actions of which the practice will be composed. The OSA has described how these objects can be articulated through the so-called configurations.

The configurations allow the analysis of mathematical practice and provide evidence of the existence or not of different types of knowledge in those who perform the mathematical task or activity, acquiring a character of epistemic configuration when referring to the network of institutional objects (and therefore correct) that are mobilized or cognitive, when referring to networks of personal objects (which may or may not be entirely correct). The relationships established between the different emergent objects in each of the configurations can be identified as semiotic functions. According to Godino et al. (2007) the notion semiotic function (SF) is the metaphorical structure that generates the correspondence between sets including: a plane of expression (initial object), a plane of content (final object) and a correspondence criterion or rule. These functions will be key to the achievement of the objective proposed in this report (Figure 2).

IDEALIZATION - MATERIALIZATION

Although the OSA does not provide an exact definition of “process”, it does enunciate some processes that are fundamental in mathematical activity; within the list are those of idealization-materialization. These try to explain how people access mathematical ideas and how they materialize them by means of sign systems, usually understood as representations of the objects being studied. The OSA proposes these processes from an ostensive - non-ostensive facet; Font and Rubio (2016) indicate that, the teaching activity embodied in the representations used to refer to a mathematical object are constituted in a concrete and ostensive part (usually accompanied by signs and a discourse that explains the idea it tries to expose) which, as a result of the idealization process is constituted in a non-ostensive object, in this case, a mathematical object that cannot be manipulated in a direct way, but only through its associated ostensive (usually known as its representations). On the other hand, to work with mathematical objects (non-ostensive), it is necessary to use ostensive representations, which come from a process of materialization (and representation).

METODOLOGY

The research follows the case study methodology (Yin, 2014). The case is that of the practices of the human group that configures a geometry didactics class, belonging to the primary education degree. A professional task (PT) is designed and implemented with a total of 250 future teachers of primary education. The subject in which the implementation was carried out is the only one that the participants have, related to the didactics of geometry within their training program. In the PT, a series of definitions and examples of similarity (given by future teachers from previous academic courses) are presented and the participants are asked to decide whether these definitions and examples are correct. Subsequently, prospective teachers are asked to give examples of geometric tasks that they would use in the classroom to work on the concept of

similarity (the analysis of the answers of this last section are the ones described in this report). The research data are the written protocols elaborated by the participants.

The notion of epistemic configuration proposed by the OSA (Godino et al., 2007) was initially used to identify the emergent connections. The configurations allow us to build a complete picture of the mathematical activity performed and the primary objects involved in PT. The data analysis was carried out in four phases following a structure like that proposed by Rodríguez - Nieto et al. (2021). The first phase consisted of the review and organization of the future teachers' responses. A first analysis was carried out to identify the mathematical practices developed by each participant. In the second phase, sequences of actions that sought to solve the proposed problems were selected from an expert perspective (carried out by the authors of this research). The third phase consisted of the construction of the epistemic configurations of the school activities proposed by the future teachers, which reflect the primary objects that emerged. Finally, in the fourth phase, possible semiotic functions were established between the primary objects of the configurations, which made it possible to form groups of semiotic functions that later served as a description of the mathematical connections. The present paper narrates the analysis process only for a professional task, allowing to visualize the emergence of the materialization connection.

RESULTS

To achieve the proposed objective of characterizing a type of extra-mathematical connection and justifying its emergence, the results obtained are grouped into three types. The first corresponds to epistemic configurations. Figure 1 below shows a fragment of the analysis of the activity proposed by group 17 (G17). The primary objects are detailed, including the concepts used, the processes and procedures carried out and the arguments provided in relation to mathematical practices.

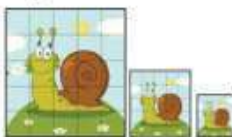
MATHEMATICAL PRACTICES		
MP01: Read the examples of similarity presented		
⋮		
MP06: Design examples of polygon similarity		
LANGUAGE		
Verbal	Graphic	Algebraic
L01: shape		If and then dice
⋮		$\triangle ABC \text{ y } \triangle DEF \angle ABC \cong$
L03: polygon		$\angle DEF, \angle BCA \cong$
⋮		$\angle EFD, \angle CAB \cong$
L05: similarity		$\angle FDE \frac{AB}{DE} = \frac{BC}{EF} =$
		$\frac{CA}{FD} \triangle ABC \sim \triangle DEF$
SITUATIONS/TASKS		DEFINITIONS
T01: Evaluate the validity of an example on similarity of geometric figures		Previews
⋮		D01: Proportionality
⋮		⋮
T06: Study and characterize the possible relationships between two photos of the same object, which have different sizes.		Emerging
		D05: Similarity

Figure 1. PT epistemic configuration by G17

The second group of results refers to the different associations and relationships established between the objects that emerged when developing each of the mathematical practices identified. This is considered as a path of semiotic functions

that, when grouped together, allow visualizing and justifying the type of connection that emerges, as well as the trajectory through which the group of future teachers accessed the mathematical object. In the case of G17, Figure 2 shows that the group goes through objects such as form (L01), the study of invariants (PC01 - PC02), the constitution of propositions as conjectures (PP03) and the use of arguments (A03) to finally access a mathematical idea (L05).

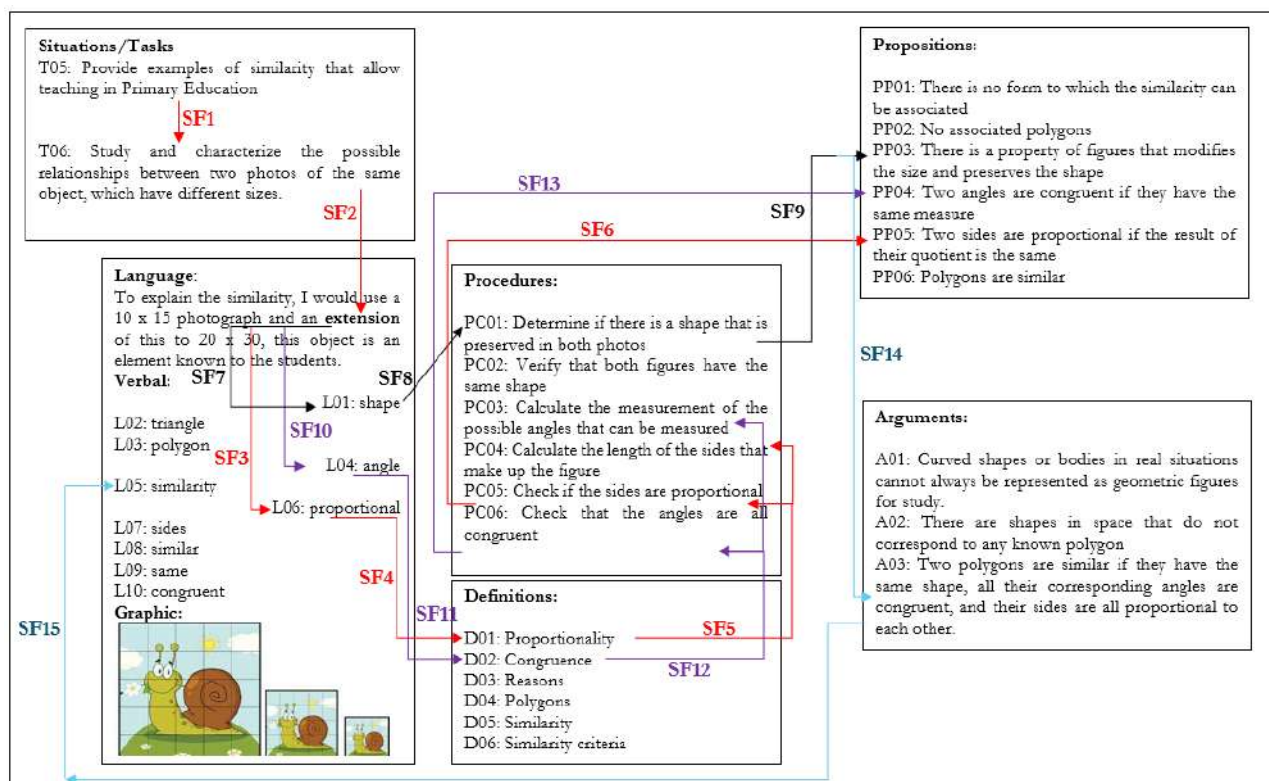


Figure 2. Epistemic configuration, semiotic functions for PT performed by G17

Finally, Figure 3 presents a detailed fragment of the instrument generated (in the case of G17) to show the emerging connections. It details the relationships established between the primary objects, the dependence that can be generated between these objects and the value of using one or another type until triggering the type of connection that emerges.

By detailing the process followed in the analysis, it is possible to evidence aspects related to the connection identified in the proposals of school activities for the teaching of similarity. If we observe the image used by group G17 as an example for the study of similarity, it is possible to indicate that there are no geometric figures that directly evidence a proportionality or congruence between their angles; even so, the proposal of the future teachers intends that their future students experience the process of idealization on such image. Expecting elementary school students to idealize the image presented in this way implies that the image (photograph in its original size) is assumed

as an ideal and explicit figure. From OSA, this image is considered as a concrete and ostensive image (since it is possible for anyone to see, touch and imagine). However, if we follow the trajectory presented in Figure 2 and the relations presented in Figure 3, it is possible to understand that the result of this idealization process is a non-ostensive object (the similarity of geometric figures), since it is a non-tangible object, which can only be reached through its associated ostensive objects.

M.P	Process	Objects	SF	Connection
MP06: Design an example of similarity that allows teaching in Primary Education	Problemization – Problem solving	T06: Study and characterize the possible relationships between two photos of the same object, which have different sizes.		MATERIALIZATION
		L06: Proportional		
		D01: Proportionality	SF1	
		PC04: Calculate the length of the sides that make up the figure	SF2 SF3	
		PC05: Check if the sides are proportional	SF4 SF5	
		PP05: Two sides are proportional if the result of their quotient is the same.	SF6 SF14	
		A03: Two polygons are similar if they have the same shape, all their corresponding angles are congruent, and their sides are all proportional to each other.	SF15	
		L05: similarity		

Figure 3. Characterization of the extra-mathematical connection obtained by G17

This detail in the analysis and the correspondence of the related primary objects allowed us to understand how the members of G17, through a process of idealization, went from an ostensive object to a non-ostensive one; implying also that, in order to manipulate this idea of similarity (in this case the non-ostensive object), some type of ostensive representation (materialization or representation) is necessary and, therefore, the photograph to be enlarged is constituted in this element that results from the materialization process and that, as a result of this research, is called extra-mathematical connection of materialization.

DISCUSSION

According to García - García (2019) research in mathematics education has the power to validate the mathematical connections that are currently known, as well as to propose the inclusion of new categories that have not yet been identified. The results of this study provide another look at the identification of connections. It makes it possible to show the existence of connections as a collection of semiotic functions that relate objects mobilized when performing any type of mathematical practice.

The materialization connection refers to the relations that allow an extra-mathematical element to materialize a mathematical idea; however, we consider the need to include in the study of connections, the process of idealization, since this process allows completing the cycle that achieves the connection. Materialization and idealization as processes have been widely studied and described by various authors, the results presented in this report regarding the processes of idealization - materialization coincide with what was raised in Font and Rubio (2016). Particularly, they agree with the suggestion of not separating ostensive objects from non-ostensive objects, in that,

although it is understood that access to mathematical objects from their representations as associated ostensive, it should not be given a character of independence of their representation to the mathematical object.

The OSA tools together with the aspects of connection theory considered have made it possible to perform different types of analysis than usual for the identification of connections. The tool that synthesizes configurations, trajectories and semiotic functions can be an ideal structure for recognizing and displaying emergent connections. We believe that the establishment and fostering of connections (both intra-mathematical and extra-mathematical) should be further studied in the context of Primary Education teacher education (Vargas et al., 2024). If these are to be addressed from the early ages, future teachers must learn to teach mathematics considering connections as a key process of school mathematical activity.

The analyses presented above correspond to a selected fragment of the study carried out on several professional tasks designed for a training course for future elementary education teachers; the theorization and description of other types of extra-mathematical connections supported by such analyses can be deepened in Vargas et. al (2024). In the case of this research report, the results are intended as an approach to the theorization of new elements on mathematical connections; it is expected that through the discussion and debate with the readers on the evidence of the way in which an extra-mathematical connection emerges in the training of future Primary Education teachers, it will be possible to contribute to the analysis and characterization of a population that has not been studied in depth in this context so far.

Methodologically, we agree with Godino et al. (2007), since the use of epistemic configurations allows the analysis of mathematical practice and provides evidence of the existence or not of different types of knowledge; in this case, the construction of the configurations and their detailed analysis in relation to mathematical processes and practices made it possible to evidence the emergence of an extra-mathematical connection. Finally, regarding the training of future primary education teachers, following de Gamboa et al. (2020) it is possible to indicate that the mathematical connections established by future teachers provide large amounts of information about the knowledge they have of the mathematical concepts they will teach; thus, the fact of promoting the use of professional tasks for their training will allow the emergence of mathematical connections and therefore, will improve the level of appropriation of mathematical ideas, providing future teachers with the necessary tools for their subsequent performance in the classroom. From the Ontosemiotic Approach, it is possible to affirm that the study of mathematical connections in the training process of future teachers will contribute to make them competent, it is they possess the necessary tools and knowledge to adequately develop the process of teaching mathematics.

Funding. This article is financed by the Regular Research Projects 2025 Competitive Fund and the Support Contest for Participation in Events 2025, of the Vice-Rector for Research and Graduate Studies of the Universidad de las Américas - UDLA.

References

- Businskas, A. M. (2008). *Conversations about connections: How secondary mathematics teachers conceptualize and contend with mathematical connections* [tesis doctoral no publicada]. Simon Fraser University, Canadá.
- de Gamboa, G. & Figueiras, L. (2014). Conexiones en el conocimiento matemático del profesor: propuesta de un modelo de análisis. En M. T. González, M. Codes, D. Arnau y T. Ortega (Eds.), *Investigación en Educación Matemática XVIII* (pp. 337 – 344) Salamanca: SEIEM.
- Font, V. & Rubio, N. (2016). Procesos en matemáticas: Una perspectiva ontosemiótica. *La matematica e la sua didattica*, 24, 97-123.
- García-García, J. G. (2019). Escenarios de exploración de conexiones matemáticas. *Números: Revista de didáctica de las matemáticas*, 100, 129-133.
- García-García, J. & Dolores-Flores, C. (2018). Intra-mathematical connections made by high school students in performing Calculus tasks. *International Journal of Mathematical Education in Science and Technology*, 49(2), 227–252. <https://doi.org/10.1080/0020739X.2017.1355994>
- Godino, J., Batanero, C. & Font, V. (2007). The Ontosemiotic approach to research in mathematics education. *ZDM—The International Journal on Mathematics Education*, 39(1–2), 127–135. <https://doi.org/10.1007/s11858-006-0004-1>
- National Council of Teachers of Mathematics [NCTM] (2000). *Principles and standards for school mathematics*. Reston, Va.: The National Council of Teachers of Mathematics (Trad. Castellana, Principios y estándares para la educación matemática. Sevilla: Sociedad Andaluza de Educación Matemática Thales, 2003).
- Rebello, N., Cui, L., Bennett, A., Zollman, D. & Ozimek, D. (2017). *Transfer of learning in problem solving in the context of mathematics and physics*. In Jonassen, D. (Ed.) *Learning to Solve Complex Scientific Problems* (1st edition, pp. 223-246). <https://doi.org/10.4324/9781315091938-10>
- Rodríguez-Nieto, C., Rodríguez-Vásquez, F., Font, V. & Morales, A. (2021). Una visión desde la red de teorías TAC-EOS sobre el papel de las conexiones matemáticas en la comprensión de la derivada. *Revemop*, 3, 1–32. <https://doi.org/10.33532/revemop.e202115>
- Vanegas, Y. & Giménez, J. (2018). Conexões extramatemáticas na formação inicial de docentes. *Estudos Avançados*, 32(94), 153–169. <https://doi.org/10.1590/s0103-40142018.3294.0012>
- Vargas Herrera, J. P., Vanegas, Y., & Giménez, J. (2024). Conexiones extra-matemáticas que establecen futurosmaestros de Educación Primaria al diseñar tareas escolares geométricas. *AIEM - Avances de investigación en educación matemática*, 25, 57-80. <https://doi.org/10.35763/aiem25.644>
- Yin, R. (2014) *Case Study Research: design and methods*. Newbury Park: Sage Publications.

EXPLORING TEACHERS' ATTEMPTS AIMING STRUGGLING AND COMPETENT STUDENTS' EQUITABLE PARTICIPATION

Maria Vasilopoulou and Chrissavgi Triantafillou

National and Kapodistrian University of Athens

This study explores conditions surrounding teachers' attempts to provide equitable participation opportunities for students they deemed as struggling or competent in learning mathematics. We focus on two teachers who conceptualize equitable participation of the target students differently. The data were participants' conceptions and their actions during specific joint participation episodes in whole-classroom discussions. We use Engeström's expanding mediational triangle for data analysis. The results illustrate that when teachers try to enhance students' confidence, it could positively impact equitable and qualitative participation chances.

INTRODUCTION

From an equity perspective, all students deserve chances to participate in classroom discourse activities (Esmonde, 2009). However, many studies report that classroom talk is often unequally distributed among students. For example, Reinholz and Shah (2018) argued that, although it might not be intentional, women and students of colour tend to be systematically given easier tasks and fewer opportunities to participate in classroom discussions. Positioning students as low achievers or struggling to understand mathematics can sometimes lead to the unintentional exclusion of these students from learning opportunities (Lack et al., 2014; Louie, 2017). Researchers emphasize examining the link between teachers' actions and their equity conceptions, as these shape in-the-moment classroom decisions (Hand, 2012). Maintaining equitable participation for diverse learners is crucial (Lack et al., 2014), starting with all students engaging in challenging tasks (Tang et al., 2017).

This study explores how two teachers strive for equitable participation for diverse learners, focusing on students, teachers labelled as *struggling* and *competent* in learning mathematics. Our research questions are: RQ1: How do teachers differentiate between competent and struggling students, conceptualize these students' participation, and enact these ideas in whole-class discussions? RQ2: What elements of equitable participation are evident in teachers' activities?

Viewing teachers and students as individuals who engage in a joint activity in the production of mathematical knowledge (Radford, 2014) we analyse teachers' actions through the lens of Engeström's (1987) expanded mediational triangle. We examine how teachers distribute participation opportunities to struggling and competent students during specific *joint participation episodes*. These episodes occur in whole-class discussions where teachers invite both struggling and competent students to participate while all are engaged in a highly demanding task, as defined by Stein &

Smith (1998). While many studies explore equitable participation of these student groups (e.g., Lack et al., 2014), few examine the conditions influencing teachers' actions when students are engaged in rich mathematical tasks.

THEORETICAL BACKGROUND

Classroom equitable participation for students with diverse learning needs

Studies often categorize students as low-attaining or slow learners or struggling and high-achieving or fast learners or competent based on their mathematical achievement (e.g., Foot & Lambert, 2011; Lack et al., 2017). Struggling students may have difficulty in solving tasks, recalling formulas, and staying focused, while competent students are typically engaged, attentive, recall quickly, and solve problems effectively (Lack et al., 2017). Roos (2019) studied students on the border of normality, finding that they do not have equal chances for active participation in classroom discourse. Skilling et al. (2015) focused on low and high achievers with low participation, highlighting the complex link between achievement and active participation.

Teachers provide participation opportunities to diverse learners in distinct ways. This may be achieved by offering enrichment opportunities for competent students and compensation opportunities for struggling students (Prediger & Buró, 2024). However, teachers may face conflicts when trying to balance equitable participation and inquiry activities (Tang et al., 2017). Equitable participation involves fairly distributing opportunities for all students to engage in learning, promoting competence, ownership, and belonging (Hand, 2012). Gazden (2001) argues that analysing equitable participation implies that “Both teachers and researchers need to monitor who participates and how, and who doesn’t and why” (p. 81).

Engeström’s activity system

Activity theory investigates human activity as a system that is object-oriented, and culturally mediated (Engeström, 1987). Engeström’s view of an activity system is represented by a triangle identifying links between mediational means (tools), the subjects, the object of the activity, and communal elements of the subjects’ activity. We focus on classroom discourse activities, in which teachers engage struggling and competent students in whole classroom discussions. We consider teachers as the *subjects* of this activity striving to address the diverse learning needs of struggling and competent students (*object*). *Tools* are the tasks and prompts that mediate teachers’ actions. The classroom environment, including its resources and expectations, constitutes the *Community*. We see *community rules* as systemic (e.g., curriculum resources) or personal (e.g., conceptions towards competent and struggling students’ participation) norms. We consider the *Division of Labour* as teachers’ responsibility for organising classroom teaching. Specifically, we consider as Division of Labour teachers’ authority to invite students to participate in classroom discussion and their responsibility to facilitate those discussions.

By taking into consideration that the distribution of opportunities to participate in mathematics discourse is an important facet of equity at the classroom level (Esmonde, 2009) we use *turn - taking organization* to explore how teachers assign work in classroom talk, an approach that is very common especially in teacher-centred classrooms. There are three typical turn-taking methods, *individual orientation* (calling individuals); *invitation to bid* (asking students to raise their hands); and *invitation to reply* (inviting the whole class to respond) (Mehan, 1979). Finally, we consider the *outcome* to be the result of the above teachers' actions in terms of the object of their activity i.e. students' equitable participation in the learning process.

METHODOLOGY

The context of the study

This study is part of a PhD project exploring how secondary mathematics teachers try to create equal participation opportunities for struggling and competent students. Ten in-service mathematics teachers (5 female and 5 male) working in different Greek public secondary schools participated. Each teacher joined 3-4 online group meetings, sharing with colleagues their challenges in creating equitable participation for students they deemed struggling or competent, and how they distinguished them. We observed and audio-recorded lessons in each participant's mathematical classroom, 4-6 times over at least two school years. Pre- and post-lesson interviews were also conducted. In pre-lesson interviews, teachers identified students they saw as struggling or competent. In post-lesson interviews, they reflected on their attempts to create equitable participation. It is worth mentioning that the participants of this project have not had any formal training in enacting equitable participation. They volunteered for this study because they wanted to find ways to enhance struggling and competent students' participation in classroom activities.

Data analysis

In terms of discussion with participants, we transcribed all interviews and group discussions, and we distinguished extracts of teachers' reports about struggling and competent students' characteristics and how they conceptualize ways of providing chances of participation for them. In terms of classroom observations, each lesson was fully transcribed. For this paper, we identified the sections of each lesson that were related to whole class discussion. We then selected joint participation episodes, which are identified as follows: a) the discussion was related to a specific highly demanding task and b) the teacher distributed participation, among other students, to *both* competent *and* struggling students. We selected the specific episodes because we wanted to explore the teachers' ways of engaging the target students in rigorous mathematical inquiry (Tang et al., 2017).

Subsequently, we coded teachers' invitations and follow-up actions according to the three turn-taking methods, i.e. individual nomination, invitation to bid, or invitation to reply. We used pseudonyms starting from 'S' (e.g., Sophie) for struggling students and

‘C’ for competent students (e.g., Chris). When other students participating in the episode are referred here as ‘StX’. Finally, we interpret teachers’ actions through the lens of Engeström’s (1987) activity system by relating data from classroom observations and discussions with the teachers.

RESULTS

In this paper, we focus on two teachers, Ms. Nelly and Ms. Elly (pseudonyms). Ms. Nelly and Ms. Elly have more than 15 years of teaching experience in Greek public schools. We chose to focus on these teachers because they teach in different schools in terms of their educational level (Ms. Nelly in Lower and Ms. Elly in Upper Secondary level) and they conceptualize competent and struggling students’ abilities and norms of participation in different ways. We present both teachers conceptions and classroom data.

Ms. Nelly’s conceptions and classroom data

Ms. Nelly categorized students as competent or struggling based on their mathematical skills and their performance in tests. Specifically, Ms. Nelly considered competent students to be those who scored “18 out of 20” and struggling students those who scored below “10 out of 20.” Moreover, competent students exhibit traits like “completing homework and consistently answering correctly in teacher’s prompts”, while struggling students often “lack basic skills such as even doing simple arithmetical calculations.” Struggling students’ participation was realized as “solving at least basic tasks and responding to easy questions”, whereas competent students “engage actively in classroom discussions, share their insights, and many times exceed teachers’ expectations.” Classroom participation, however, is not a definitive marker of a student’s competence, since there were cases where “high-performing students remained quiet during classroom discussions but did very well in tests.”

The joint participation episode: Ms. Nelly’s Grade 9 class has 26 students, including 7 struggling and 9 competent students. The specific episode lasted for 9 minutes. In this episode, the teacher distributed participation to five struggling students (Stacy, Stella, Stathis, Savina, and Sandy) and four competent students (Cas, Chris, Clara, Colin). Ms. Nelly’s task was “Explore the number of solutions of the equation $x + y = 10$.” The task aimed for students to understand that this equation has infinitely many solutions, which are collinear points of the function $y = 10 - x$.

In the first part, Ms. Nelly invited only struggling students to the conversation floor. Specifically, Ms. Nelly started with individual nomination, asking Stacy, “How many solutions do we have Stacy?” Stacy answered: “We have many solutions.” Ms. Nelly continued with follow-up questions, asking Stacy to indicate specific solutions: “I want you to give me a couple of solutions, one for x and one for y .” Stacy replied: “5 plus 5.” Then, Ms. Nelly continued: “Stacy, look at the equation again; it has different letters for each variable.” Stacy changed her initial response to “6 plus 4.” Nelly continued with individual nomination to Stella, Stathis, Savina, and Sandy who provided

solutions with natural numbers. For example, Savina said, “9 plus 1,” and Sandy, “7 plus 3.”

In the second part, Ms. Nelly only invited competent students to the conversation floor. These students initially provided solutions with integers and rational numbers. They also attempted, unsuccessfully to provide solutions with irrational numbers. Specifically, Ms. Nelly invited Cas with individual nomination, and he responded “-10 plus 20.” Then, Nelly invited all students to reply: “I have noticed that you restrict yourselves to integers. Are there any other solutions?” Chris and Carla took up the conversation floor and responded. Chris said, “It could be decimal numbers like 9.9 plus 0.1,” and Carla said, “ $\sqrt{25}$ plus $\sqrt{25}$.” Then, Ms. Nelly responded to Carla: “Carla, is $\sqrt{25}$ an irrational number? What is the square root of 25?” Carla replied, “It’s 5,” and admitted that it is not an irrational number. Ms. Nelly then provided a pair of irrational solutions herself: “ $x = 5 + \sqrt{2}$ and $y = 5 - \sqrt{2}$.”

Ms. Nelly summarized the different pairs of solutions provided before and said: “This equation can take many different solutions; thus, the solutions are...” [invitation to reply] a StX added, “Infinite many solutions.” Then, two competent students (Colin and Chris) continued as follows: Colin said, “Thus, it is an algebraic identity.” Chris asked, “Does the fact that it has infinite many solutions make it an identity?” Ms. Nelly intervened, “Good question Chris! What is an algebraic identity?” Chris replied, “No matter what value we put on the variables, it becomes true.” Ms. Nelly responded, “Chris, for any value for x and y, will their sum be 10?” Chris answered “No.” Ms. Nelly concluded, “Well-done Chris, this equation has infinite many solutions, but it is not an algebraic identity. It is a function in the form of $y = 10 - x$ ”.

Ms. Elly’s conceptions and classroom data

Ms. Elly differentiates competent from struggling students by their confidence in solving open-ended tasks. Struggling students often lack confidence, remaining reserved and hesitant to engage.

“They do not feel confident in doing mathematics. You see them being quiet and shy; they do not want to be exposed, but I always nudge them and try to find ways to get them involved and strengthen their confidence.”

Ms. Elly added that she actively supports struggling students’ participation through desk-based responses but never asks them to go to the blackboard “to reduce their anxiety.” She views participation as involving participation in classroom discussions and attentiveness, asserting that competent students consistently engage, while struggling students require encouragement to become involved. She admitted that “she never had competent students not participating in classroom discussions” and “peer supporting enhance students’ confidence since it is a way of their knowledge verification.”

Joint participation episode: Ms. Elly’s Grade 10 class has 27 students, including 6 struggling and 7 competent students. Ms. Elly’s task was “Finding out the recursive

sequence in the row elements of Pascal's triangle." In this task, students had to understand the recursive sequence of the row elements of the well-known triangle. In the beginning, Ms. Elly wrote the first 3 rows on the board and invited StX to complete the 4th row i.e. 1 3 3 1. Then, Ms. Elly invited Sophie with individual nomination to identify the 5th row elements, i.e. 1 4 6 4 1 (Fig. 1), "Come on Sophie." Sophie responded: "1, 4, 4, 4, 1." Ms. Elly asked, "How did you come to this result, Sophie? The 4 [on the 3rd column] is not correct. Can you think of a different number?" Then StX shouted 'add,' obviously to help Sophie. Sophie responded, "Do we add? ... Is it five?". Ms. Elly responded, "What did you add to find five?" Sophie seemed confused: "2 and 3?"

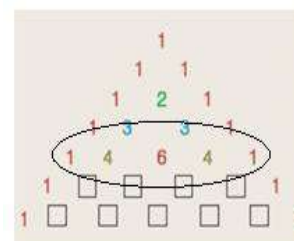


Figure 1: The requested line of Pascal's triangle

Ms. Elly recognized the difficulty Sophie faced in discerning the underlying rule. To address this, Ms. Elly asked the entire class to reply, "Yes, adding is correct, but which elements do we add?" Carol then contributed by stating, "Working like in a triangle". Ms. Elly decided to invite Carol to assist Sophie, so she said, "Carol, I don't understand." Carol explained, "I added, I put 1, then I put 1 plus 3, which makes 4, 3 plus 3 which makes 6, 3 plus 1, which makes 4, and 1 again at the end". Subsequently, Ms. Elly redirected the conversation to Sophie, asking, "Sophie, do you understand what Carol is saying? Can you tell me what the next [6th] row is?" Sophie did not initially respond, so Ms. Elly prompted Carol to repeat her explanation. Then, Ms. Elly directed the question to Sophie again, "So, Sophie, can you tell me the next row?" Sophie promptly and accurately replied, "1, 1 plus 4 makes 5, 4 plus 6 makes 10, 6 plus 4 makes 10 again, and 1." The teacher then wrote Sophie's answer on the board, confirming the accuracy of Sophie's response, and other students (StX) provided the next 5 rows of Pascal triangle.

Analysing teachers' actions through the lens of Engeström's activity system

Two mathematics teachers, Ms. Nelly and Ms. Elly, engaged in joint participation episodes with their students. The object of Ms. Nelly's activity was identifying the infinite numbers of solutions to the equation $x + y = 10$. The object of Ms. Elly's activity was exploring recursive sequences within Pascal's Triangle. Both teachers utilized highly demanding tasks, prompts, and specific turn-taking strategies as mediating tools. We recognize the classroom as a community, influenced by systemic (curriculum) guidelines and personal rules (teachers' conceptions). While both teachers valued student participation in discussions, their definitions of competent and struggling students differed. Ms. Nelly relied on academic achievement, while Ms. Elly emphasized student confidence in tackling challenging tasks. These differing perspectives influenced their division of labour actions. Ms. Nelly employed a linear approach, inviting struggling students first for simpler solutions and then engaging competent students for more complex ones. In contrast, Ms. Elly used a circular

pattern, alternating between struggling and competent students, and fostering peer support. Also, Ms. Nelly occasionally provided answers when students could not respond to a question, while Ms. Elly empowering students to take ownership of their learning and build their confidence in tackling high demanding tasks. The outcome of two teachers' activity had different qualitative characteristics. In Ms. Nelly's classroom activity, even though there were nine participant students, only one competent student seems to grasp the central task's idea. On the other hand, in Ms. Elly's activity a struggling student developed a significant arithmetic generalization with a competent student's help. Thus, equitable participation, in relation to development of struggling and competent students' mathematical competence and knowledge ownership, seems to be accomplished only in Ms. Elly's class.

CONCLUDING REMARKS

This study, though limited by a small sample size, provides in-depth insights into how teachers' conceptions about struggling and competent students' abilities impact their enactment of equitable participation chances, as other researchers also mention (e.g., Hand, 2012). Specifically, while the study outcomes argue that both teachers aimed to engage struggling and competent students in highly demanding inquiry whole class discussions (e.g., Tang et al., 2017), they differed in creating fair participation opportunities. When the teacher was restricting struggling students' participation only to the easy part of the highly demanding task, this resulted in limiting students' equitable participation aligning with other studies' findings (e.g., Reinholz & Shah, 2018), and in unintentional exclusion (Louis, 2017). Furthermore, the results suggest that when the teacher tries to enhance students' confidence by engaging them in highly demanding tasks, insisting on succeeding in responding and/or providing them chances to support their peers, could result in equitable participation. Finally, the fair distribution of opportunities for participation seems not to be related to the number of participants but to the quality of their participation.

REFERENCES

- Cazden, C.B. (2001). *Classroom discourse: The language of teaching and learning* (2nd ed.). Ports- mouth, NH: Heinemann.
- Engeström, Y. (1987). *Learning by Expanding: An activity theoretical approach to developmental research*. Helsinki: Orienta-Konsultit.
- Ernest, J. B., Reinholz, D. L., Shah, N. (2019). Hidden competence: Women's mathematical participation in public and private classroom spaces. *Educational Studies in Mathematics*, 102, 153-172. <https://doi.org/10.1007/s10649-019-09910-w>
- Esmonde, I. (2009). Ideas and identities: Supporting equity in cooperative mathematics learning. *Review of Educational Research*, 79(2), 1008–1043. <https://doi.org/10.3102/0034654309332562>

- Foote, M. Q., & Lambert, R. (2011). I have a solution to share: Learning through equitable participation in a mathematics classroom. *Canadian Journal of Science, Mathematics and Technology Education*, 11(3), 247-260. <https://doi.org/10.1080/14926156.2011.595882>
- Hand, V. (2012). Seeing culture and power in mathematical learning: Toward a model of equitable instruction. *Educational Studies in Mathematics*, 80(1), 233-247. <https://doi.org/10.1007/s10649-012-9387-9>
- Lack, B., Swars, S. L., & Meyers, B. (2014). Low-and high-achieving sixth-grade students' access to participation during mathematics discourse. *The elementary school journal*, 115(1), 97-123. <https://doi.org/10.1086/676947>
- Louie, N. L. (2017). The culture of exclusion in mathematics education and its persistence in equity-oriented teaching. *Journal for Research in Mathematics Education*, 48(5), 488-519. <https://doi.org/10.5951/jresmetheduc.48.5.0488>
- Mehan, H. (1979). *Learning lessons: Social organization in the classroom*. Harvard University Press.
- Prediger, S., & Buró, R. (2024). Fifty ways to work with students' diverse abilities? A video study on inclusive teaching practices in secondary mathematics classrooms. *International Journal of Inclusive Education*, 28(2), 124-143. <https://doi.org/10.1080/13603116.2021.1925361>
- Radford, L. (2014). On teachers and students: An ethical cultural-historical perspective. In P. Liljedahl, C. Nicol, S. Oesterle, & D. Allan (Eds.), *Proceedings of the Joint Meeting of PME 38 and PME-NA 36* (Vol. 1, pp. 1-20). Vancouver: PME.
- Reinholz, D. L., & Shah, N. (2018). Equity analytics: A methodological approach for quantifying participation patterns in mathematics classroom discourse. *Journal for Research in Mathematics Education*, 49(2), 140-177. <https://doi.org/10.5951/jresmetheduc.49.2.0140>
- Roos, H. (2019). Challenges at the border of normality: Students in special educational needs in an inclusive mathematics classroom. In *Mathematics Education and society, 10th International Conference*, Hyderabad, India, Jan 28th-Feb2nd, 2019 (pp. 928-940). Mathematics Education and Society.
- Skilling, K., Bobis, J., & Martin, A. (2015). The engagement of students with high and low achievement levels in mathematics. In *Proceedings of the 39th Psychology of Mathematics Education conference* (Vol. 4, pp. 185-192). Psychology of Mathematics Education (PME).
- Stein, M. K., & Smith, M. S. (1998). Mathematical tasks as a framework for reflection: From research to practice. *Mathematics teaching in the middle school*, 3(4), 268-275. <https://doi.org/10.5951/MTMS.3.4.0268>
- Tang, G., El Turkey, H., Cilli-Turner, E., Savic, M., Karakok, G., & Plaxco, D. (2017). Inquiry as an entry point to equity in the classroom. *International Journal of Mathematical Education in Science and Technology*, 48(sup1), S4-S15. <https://doi.org/10.1080/0020739X.2017.1352045>

MATHEMATICIANS' VIEWS ON THE SUBJECT-SPECIFIC EDUCATION OF PROSPECTIVE TEACHERS

Birke-Johanna Weber¹, Anika Dreher², Aiso Heinze³, & Daniel Sommerhoff³

¹Kiel University, Germany; ²Freiburg University of Education, Germany;

³Leibniz Institute for Science and Mathematics Education, Kiel, Germany

Teachers need to understand the underlying ideas of the mathematical contents they teach in school. Thus, prospective upper secondary school teachers are often required to take several academic mathematics courses during their studies at university. These courses are often designed by mathematicians. However, it is so far unclear to what extent these mathematicians consider school-related content knowledge (SRCK) as a relevant part of teachers' subject-specific knowledge. The results of our online survey with 424 German mathematics lecturers indicate that they consider SRCK as a relevant part of teachers' knowledge. Furthermore, they mostly believe that academic mathematical content knowledge is insufficient to acquire SRCK, and that SRCK needs specific learning opportunities.

INTRODUCTION

In many countries, university education for upper secondary school teachers traditionally includes many academic mathematics courses (Tatto et al., 2010). These courses are usually designed by mathematicians who thus shape a significant part of mathematics teachers' education (Leikin et al., 2018). Furthermore, the courses are often not only attended by prospective teachers but also by mathematics students, and explicit connections to school mathematics are not systematically made. One (implicit) reason for this may be the so-called *intellectual trickle-down hypothesis* (Wu, 2011): If prospective teachers are able to understand abstract academic mathematics, they should automatically be able to relate their knowledge of academic mathematics to school mathematics.

However, for more than 100 years, it has been argued that (prospective) teachers may not be able to make these connections on their own. Felix Klein (1908) suggested that they have difficulties making connections between school mathematics and academic mathematics. As a result, they may not be able to use their academic mathematical knowledge for teaching. In line with Klein's assumption, qualitative studies show that (prospective) teachers perceive academic mathematics as having little or no relevance for teaching (e.g., Even, 2011; Zazkis & Leikin, 2010). Furthermore, the quantitative study by Hoth et al. (2020) showed that prospective teachers, in fact, do not automatically acquire knowledge about connections between school mathematics and academic mathematics, so-called *school-related content knowledge* (SRCK, Dreher et al., 2018) when they acquire academic mathematical content knowledge. In sum, the existing empirical evidence does not support the intellectual trickle-down hypothesis.

Several initiatives aim to strengthen the relevance of academic mathematics courses for prospective upper secondary school teachers (for an overview see, e.g., Wasserman et al., 2023). In this context, mathematicians play a central role. They are de facto mathematics teachers' educators, even if they might not identify as such (Hoffmann & Even, 2024; Leikin et al., 2018), and they shape a crucial part of learning opportunities in prospective teachers' studies. An open question is to what extent mathematicians consider school-related content knowledge as a relevant part of teachers' subject-specific professional knowledge, and, if so, whether they think SRCK needs explicit learning opportunities. Mathematicians' views about these questions are likely to determine the learning opportunities they provide in their courses or are willing to provide in the future. In the following, we present a study from Germany that examines whether mathematicians believe that academic content knowledge is sufficient for prospective teachers' subject-specific education and – if not – where mathematicians would like to locate learning opportunities for school-related content knowledge within teacher education programs.

THEORETICAL BACKGROUND

There is a broad consensus that prospective secondary mathematics teachers at upper secondary school need academic content knowledge. However, there is an ongoing debate about what kind of subject-specific knowledge mathematics teachers need and how it impacts on student learning. This is related to the question of what knowledge areas should be addressed in academic mathematics courses for prospective teachers.

Teachers' subject-specific knowledge

Monk (1994) found that the number of academic mathematics courses taken by teachers did not generally affect their (later) students' achievement. Instead, the type and quality of knowledge actually acquired by a teacher appear to be more important. Pedagogical content knowledge (PCK) has been found to be a stronger predictor of student achievement than academic mathematical content knowledge (CK), but CK is considered to be a prerequisite for PCK (e.g., Krauss et al., 2008). Still, recent research not only distinguishes between CK and PCK, but also assumes that teachers need a profession-specific knowledge about connections between school mathematics and academic mathematics (e.g., Ball et al., 2008; Dreher et al., 2018; Wasserman et al., 2023), which we refer to as school-related content knowledge (SRCK).

SRCK is a profession-specific content knowledge for secondary teachers about connections between school mathematics and academic mathematics. It is conceptualized in three facets (Dreher et al., 2018): Teachers need to be able to make connections from school mathematics to academic mathematics (bottom-up), for example, when evaluating an unexpected student statement against the background of academic mathematics. They also need knowledge about connections from academic mathematics to school mathematics (top-down), for example, when teachers want to make an academic mathematical aspect (e.g., the construction of the real numbers) accessible within school mathematics. Knowledge about top-down connections also

concerns mathematical methods. For example, teachers need to consider how they can present defining (a central activity in academic mathematics) in their lessons in an intellectually honest way so that their students learn what characterizes a “good” definition. The third facet, curricular knowledge, includes knowledge about the structure of school mathematics and the underlying reasons from an academic perspective. Teachers need this kind of knowledge, for example, when deciding which ideas they use to explain a mathematical concept in a particular class, taking into account which concepts their students have learned earlier and which ideas might prepare them for concepts that will be introduced later.

In summary, upper secondary school mathematics teachers need substantial academic content knowledge which is a prerequisite to develop SRCK. Still, there seems to be no strong intellectual trickle-down effect, so academic mathematical content knowledge does not seem to be sufficient for the acquisition of SRCK (Hoth et al., 2020), so that SRCK needs explicit learning opportunities. However, it seems that such learning opportunities are not implemented systematically in teacher education courses – although there is a growing number of initiatives to strengthen academic mathematics courses for prospective teachers (Wasserman et al., 2023). One reason may be that academic mathematics courses are often designed by mathematicians who are usually not familiar with the current discourse in mathematics didactics. To strengthen academic mathematics courses and systematically implement school-related learning opportunities, it is therefore crucial to understand the perspectives mathematicians hold on their mathematics courses for prospective teachers.

Mathematicians’ views on mathematics courses for teachers

A few studies have examined mathematicians’ perspectives on the relevance of academic mathematics courses for secondary teachers. For example, Hoffmann and Even (2024) found three main questions that mathematicians considered important for prospective teachers: What is mathematics? How is mathematics done? Why engage with mathematics? In sum, they found that mathematicians saw knowledge about the essence of mathematics as a main objective in their courses for prospective teachers. However, the interviewed mathematicians did not suggest how teachers could integrate this knowledge into their teaching. It is unclear whether they felt unqualified to do so (which was mentioned by some participants) or whether they expected an automatic trickle-down effect to happen. Furthermore, Leikin et al. (2018) found that mathematicians could talk a lot about the relevance of academic mathematics for teachers on a meta-level, but did not give many concrete examples. If they gave concrete examples, they mostly referred to extracurricular content that teachers could present to high-achieving students. Another finding of Leikin et al. (2018) is that mathematicians do not seem to pay particular attention to the fact that they are educating prospective teachers when they design their lectures. Based on these qualitative findings, one could assume that school-related learning opportunities do not play a central role in many academic mathematics courses. It is possible that

mathematicians think that it is sufficient to teach academic mathematics without explicitly addressing school-related content knowledge. But it is also possible that they see others responsible for addressing SRCK – for example courses for mathematics didactics or courses during practical teacher training phases after graduation. To be able to clarify this question and identify potentials to systematically strengthen the subject-specific education of prospective mathematics teachers, it is necessary to better understand the views of mathematicians as they shape a central part of learning opportunities for prospective teachers.

Research question

The research questions arising from the theoretical background presented above are:

RQ1: To what extent do mathematicians consider SRCK as a relevant part of the subject-specific education of prospective upper secondary school teachers?

RQ2: Do mathematicians think that SRCK needs specific learning opportunities and that academic mathematics courses should be adapted?

As our study was conducted in Germany, we refer to “upper secondary school teachers” as teachers at the academic track of secondary school (Gymnasium) teaching at lower and upper secondary level. Our expectation is that most mathematicians see academic mathematical content knowledge as well as SRCK as relevant for teachers. However, following the reported qualitative results on mathematicians’ views and the rather traditional canon of teacher education at German universities, we also expect them to follow the intellectual trickle-down hypothesis and consider academic mathematical content knowledge to be sufficient to acquire SRCK. If mathematicians believe that further learning opportunities are necessary to develop SRCK, we expect them to not see academic mathematics courses but mathematics didactics courses or the practical teacher training phase after graduation (a compulsory part of teacher education in Germany) as the best places to locate these learning opportunities.

METHOD

Based on the information given on the websites of all German universities offering degree programs for prospective upper secondary mathematics teachers, we identified 1218 mathematicians (professors, adjunct professors, junior professors, and habilitated lecturers called “Privatdozent”) who were potentially involved in academic mathematics courses for prospective teachers. We invited them to participate in our online survey in the summer of 2021. 520 mathematicians from all over Germany participated in the survey, and we included $N = 424$ datasets in the analysis. Individuals were excluded from the analysis if they had not taught a course for prospective teachers in the past five years, did not answer any question in the survey, or identified as researcher in mathematics didactics. Most participants (77.4%) held a professorship.

To ensure a shared understanding of different subject-specific knowledge areas, we introduced the following four knowledge areas at the beginning of the survey (including examples for illustration): (1) *Academic content knowledge*: The kind of

knowledge taught in academic mathematics courses at university. It refers to mathematics in a scientific context. (2) *School mathematical knowledge*: The kind of knowledge taught in school. Good students have this knowledge when they graduate from school. (3) *School-related content knowledge*: Knowledge about connections between school mathematics and academic mathematics as well as knowledge about the structure of school mathematics. (4) *Pedagogical content knowledge*: Knowledge about mathematical teaching and learning processes taught in courses for mathematics didactics at university.

The survey was structured in three parts: (1) *Relevance of the subject-specific education for prospective mathematics teachers*: We asked the participants to rate the relevance of different areas of knowledge (school mathematical knowledge, academic mathematical content knowledge (basic vs. extended), school-related content knowledge) for teachers. (2) *Intellectual trickle-down hypothesis*: The mathematicians were asked to state whether they believe that academic content knowledge is sufficient for the development of SRCK or whether SRCK needs specific learning opportunities. Participants who rejected the hypothesis were asked where they would locate further learning opportunities for SRCK in teacher training programs. (3) *Ideas for improving the subject-specific education for prospective teachers*: We asked the participants to indicate what changes they would make to the weighting of learning opportunities for academic content knowledge and learning opportunities for SRCK at their university within the framework of the available credit points. 34 participants were excluded from the analysis of the third question because they indicated that they did not know the weighting at their university and, therefore, could not answer it. Finally, we asked the participants what they see as an ideal distribution of the three subject-specific knowledge areas (academic, school-related, and pedagogical content knowledge).

RESULTS

Our participants rated the relevance of all knowledge areas surveyed (school mathematical knowledge, academic content knowledge, and school-related content knowledge) as high or very high for teachers (see Tab. 1).

Knowledge area	<i>N</i>	<i>M (SD)</i>
School mathematical knowledge	407	5.38 (1.10)
School-related content knowledge	409	5.35 (1.02)
Academic content knowledge		
- Basic (introductory courses: linear algebra I/II, real analysis I/II, stochastics)	410	5.49 (0.87)
- Extended (beyond introductory courses)	410	4.45 (1.24)

Table 1: Assessed relevance of different knowledge areas for teachers. The scale ranged from “0 – very irrelevant” to “6 – very relevant”.

Furthermore, we found that a majority of 63.2% did not support the intellectual trickle-down hypothesis. Contrary to our expectations, only 36.8% of our sample believed that academic content knowledge was sufficient to acquire SRCK. Participants who disagreed with the hypothesis and believed that SRCK needed specific learning opportunities were asked where these learning opportunities should be located. A minority of 17.1% stated that SRCK learning opportunities should be located at the postgraduate practical teacher training phase (a compulsory part of teacher education in Germany), 37.8% preferred university courses for mathematics didactics, and most participants (45.2%) named academic mathematics courses as the preferred place.

When asked if and how they would change the weighting of credit points between courses for academic content knowledge and courses for SRCK, about half of the sample (44.7%) reported that they would not change the weighting. 32.1% said they would shift the weighting in favor of academic content knowledge at the expense of SRCK, while 23.2% preferred a shift toward SRCK. According to our participants, for an ideal distribution, on average, half of the available credit points would be allocated to academic content knowledge and the other half to SRCK and PCK in roughly equal parts (see Tab. 2). However, there was quite a bit of variation, suggesting that our participants differed substantially in their views about the ideal composition of subject-specific education for prospective teachers.

	Academic content knowledge	School-related content knowledge	Pedagogical content knowledge
<i>M (SD)</i>	<i>54.1 (17.7)</i>	<i>21.8 (11.8)</i>	<i>24.1 (10.7)</i>

Table 2: Percentage of credit points participants ($N = 344$) would ideally assign to academic, school-related, and pedagogical content knowledge.

DISCUSSION AND OUTLOOK

The results of our survey revealed a more progressive picture of mathematicians' views on the subject-specific education of prospective secondary teachers than we expected based on the rather traditional canon in German university teacher education programs. First, we found that the participants not only rated academic mathematical content knowledge and school mathematical knowledge as relevant for teachers but also school-related content knowledge. They would, on average, attribute as many credit points to courses for SRCK as they would to courses for pedagogical content knowledge. Second, and more surprisingly, most mathematicians in our sample rejected the intellectual trickle-down hypothesis that academic mathematical content knowledge is sufficient for acquiring SRCK (Wu, 2011), meaning they believed that SRCK needs specific learning opportunities. Third, also contrary to our expectations, more than 40% of the participants see academic mathematics courses as the place where these learning opportunities should be implemented rather than didactic courses or a postgraduate practical teacher training phase. This may mean that they see themselves as having a responsibility to support prospective teachers in acquiring

SRCK. Therefore, our results suggest that recent initiatives (e.g., Wasserman et al., 2023) such as the integration of capstone courses or school-related mathematical problems could potentially be implemented more in the breadth of university teaching in the future. This raises the question of how much knowledge about connections between school mathematics and academic mathematics mathematicians themselves need to have in order to teach SRCK in their own courses (Lai et al., 2019).

Of course, our study has some limitations. First, our study was conducted in Germany, meaning the results must be interpreted in the context of the organization of teacher education programs in Germany. Nevertheless, the large number of academic mathematics courses typical for teacher education in Germany can also be found in many other countries (Tatto et al., 2010). Second, our study focused on the subject-specific education of prospective teachers at the academic track secondary school (Gymnasium). We expect the results would be different if mathematicians were asked for their opinion on the education of other teacher training programs (e.g., for lower secondary school). These programs usually have fewer academic mathematics courses and a greater focus on didactics and pedagogy. Third, our study does not provide information on whether and how SRCK is currently implemented at the surveyed universities. We focused on mathematicians' views because they are likely to influence their teaching, and we did not survey the curriculum currently implemented at their universities. Half of our sample suggested that no changes should be made to the current weighting of courses for academic content knowledge and SRCK. This could mean that there are already learning opportunities for SRCK that mathematicians value or that there are no learning opportunities for SRCK what other mathematicians who believe in the intellectual trickle-down hypothesis might prefer (see Weber et al., accepted). Presumably, both variants are represented in our sample.

Summarizing, we would like to emphasize that about half of our sample suggested changes to the current weighting of credit points regarding academic and school-related content knowledge at their university. This suggests a relatively high potential for change within the curriculum of academic mathematics courses for prospective secondary teachers. To better assess this potential, the next step is to identify mathematicians' arguments for their proposed changes. Understanding the views mathematicians hold on the subject-specific education of prospective secondary teachers can help to identify starting points for the further development of teacher training programs and to systematically integrate school-related learning opportunities in university courses for prospective teachers.

References

- Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, 59, 389–407.
- Dreher, A., Lindmeier, A., Heinze, A. & Niemand, C. (2018). What kind of content knowledge do secondary mathematics teachers need? *Journal für Mathematik-Didaktik*, 39, 319–341.

- Even, R. (2011). The relevance of advanced mathematics studies to expertise in secondary school mathematics. *ZDM – Mathematics Education*, 43, 941–950.
- Hoffmann, A., & Even, R. (2024). What do mathematicians wish to teach teachers about the discipline of mathematics? *Journal of Mathematics Teacher Education*, 27, 479–497.
- Hoth, J., Jeschke, C., Dreher, A., Lindmeier, A. & Heinze, A. (2020). Ist akademisches Fachwissen hinreichend für den Erwerb eines berufsspezifischen Fachwissens im Lehramtsstudium? Eine Untersuchung der Trickle-down-Annahme [Is academic content knowledge sufficient for the acquisition of subject-specific professional knowledge during university teacher education? An investigation of the trickle-down hypothesis]. *Journal für Mathematik-Didaktik*, 41, 329–356.
- Klein, F. (1908). *Elementarmathematik vom höheren Standpunkt aus. Arithmetik, Algebra, Analysis* [Elementary mathematics from a higher standpoint. Arithmetic, algebra, analysis]. Springer.
- Krauss, S., Baumert, J., & Blum, W. (2008). Secondary mathematics teachers' pedagogical content knowledge and content knowledge. *ZDM – Mathematics Education*, 40, 873–892.
- Lai, Y., Strayer, J., Lischka, A., Anhalt, C., Quinn, C., & Reed, S. (2019). Theoretical report: A framework for examining prospective teachers' use of mathematical knowledge for teaching in mathematics courses. In A. Weinberg, D. Moore-Russo, H. Soto & M. Wawro (eds.), *Proceedings of the 22nd annual conference on research in undergraduate mathematics education* (pp. 765–733). SIGMAA.
- Leikin, R., Zazkis, R., & Meller, M. (2018). Research mathematicians as teacher educators: focusing on mathematics for secondary mathematics teachers. *Journal of Mathematics Teacher Education*, 21, 451–473.
- Monk, D. H. (1994). Subject area preparation of secondary mathematics and science teachers and student achievement. *Economics of Education Review*, 13, 124–145.
- Tatto, M. T., Lerman, S., & Novotna, J. (2010). The organization of the mathematics preparation and development of teachers: A report from the ICMI Study 15. *Journal of Mathematics Teacher Education*, 13, 313–324.
- Wasserman, N. H., Buchbinder, O., & Buchholtz, N. (2023). Making university mathematics matter for secondary teacher preparation. *ZDM – Mathematics Education*, 55, 719–736.
- Weber, B.-J., Sommerhoff, D., Dreher, A. & Heinze, A. (accepted). Welches fachliche Wissen sollten angehende Gymnasiallehrkräfte erwerben? Einschätzungen von Hochschullehrenden [What subject-specific knowledge should prospective secondary school teachers acquire? Perspectives of lecturers]. *Mitteilungen der DMV*.
- Wu, H.-H. (2011). The mis-education of mathematics teachers. *Notices of the AMS*, 58, 372–384.
- Zazkis, R., & Leikin, R. (2010). Advanced Mathematical Knowledge in Teaching Practice: Perceptions of Secondary Mathematics Teachers. *Mathematical Thinking and Learning*, 12, 263–281.

ECOLOGICAL CONCEPTUALIZATIONS IN MATHEMATICS TEACHER EDUCATION AND TEACHER NOTICING

Tracy L. Weston

Middlebury College

To understand the possibility and implications of recent proposals for ecological models of teacher noticing (Jazby et al., 2023; Scheiner, 2021), mathematics teacher educators and researchers must first understand what is meant by ecological. This study provides a systematic review of the literature, and investigated the (a) language used to describe and (b) theories used to conceptualize ecological approaches in mathematics teacher education. Ecological conceptualizations in mathematics teacher education rarely used ecology. Most often, “ecological” was used as an imprecise descriptor without any explanatory power, and the majority of ecological conceptualizations contained elements that were inconsistent with ecology. Implications for teacher noticing are discussed.

INTRODUCTION

Ecological frameworks, models, and approaches have been proposed in various aspects of teacher education, including student teaching supervision (Buchanan, 2020), teacher professional development (Ehrenfeld, 2022), and mathematics teacher noticing (Scheiner, 2021). Recent teacher noticing conceptualizations have invited expanding thinking beyond cognitive theories of noticing (e.g., attend, interpret, respond, Jacobs et al., 2010) to alternate models that “present an ecological account of noticing which shifts attention towards the teaching environments so that teacher-level psychological constructs are not the only factors that contribute to teacher noticing” (Jazby et al., 2023, p. 647). As someone who teaches courses in mathematics pedagogy, science pedagogy, and environmental education, and also researches teacher noticing, I am particularly interested in invitations to consider an ecological framing of teacher noticing (Jazby et al., 2023; Scheiner, 2021).

In the United States, *ecology* is typically defined as the discipline of studying organisms, their interactions, and their interactions with the environment—biotic and abiotic (D. Allen, personal communication, June 25, 2024). Shifting from a noun to an adjective (*ecologic*, *ecological*) or adverb (*ecologically*) maintains a similar definition tied to ecology, although colloquially these words have taken on a more flexible use and are attached to a range of activities, issues, fields, and marketing campaigns, from eco car washes to ecological economics. Jacobs and Spangler (2017) indicated concern about the lack of a coherent use of the term *noticing* and suggested that if future researchers are not clear, the term could eventually be all encompassing and thus lack meaning. Considering the recent incorporation of *ecological* with teacher noticing, I share a similar desire for clarity and specificity about what is meant by *ecological*.

Given that teacher noticing researchers have used varied and unclear language (Weston & Amador, 2023), it follows that *ecological teacher noticing*, or *noticing in ecological ways*, could likewise benefit from explicit conceptualization and definition. Therefore, this project investigated the use of *ecological* across mathematics teacher education research, in order to apply those findings to teacher noticing. Just as there is a “general consensus that teacher noticing differs from everyday noticing” (Dindyal et al., 2021, p. 2), it is important to investigate what *ecological* means within the context of mathematics education research, so that resulting theories for teacher noticing—and other areas—are robust.

RELATED LITERATURE

An early example of *ecological* appearing in education research is Bronfenbrenner’s (1979) ecological systems theory, or bioecological approach. Bronfenbrenner conceptualized human development as a series of complex interactions that influenced an individual based on their development within a series of nested, interconnected relationships, institutions, and systems. His model indicates a range of environmental influences including the individual, family and peer relationships, schooling, the media, and broader political, social, and cultural contexts. Gibson’s (1979) work introduced an ecological approach to perception and emphasized the importance of the environment. He was “concerned...with things at the ecological level, with the habitat of animals and men, because we all behave with respect to things we can look at and feel, or smell and taste, and events we can listen to” (p. 5). Recently, Gibson’s (1979) work on perception has been used to conceptualize noticing beyond a cognitive activity that solely resides in the mind of a teacher. Jazby et al.’s (2023) work presents an ecological framing for teacher noticing that proposes a perception-action cycle in which the teacher and classroom interact, rather than noticing being solely under the control of the teacher. Scheiner’s (2021) dynamic model of teacher noticing includes an embodied-ecological aspect to describe the “perceptual activity and exploration” (p. 91) that a teacher actively participates in and constructs through their interaction in and with the classroom environment.

RESEARCH AIM AND QUESTIONS

The aim of this research is to seriously consider ecology, how it is (mis)applied in mathematics teacher education research, and opportunities to use ecology in “ecological approaches.” The specific research questions are: (1) How have researchers described and conceptualized ecological approaches in mathematics teacher education? (2) What aspects of their self-described ecological approaches are consistent and inconsistent with ecology? (3) What insights can we gain from the approaches that used ecology? Implications for teacher noticing will also be discussed.

METHOD

This synthesis traced ecological theories and frameworks found in mathematics teacher education research literature. Since the aim of this study was to make sense of

ecological theories to then apply that understanding to emergent research on teacher noticing, publications included, but were not limited to, research on mathematics teacher noticing. To identify as many relevant publications as possible, a search was conducted for “teacher education AND math* AND ecolog*.” The second wildcard found publications that used either *ecology* or *ecological*. This search was conducted in April 2024 within titles, abstracts, and keywords in ERIC, PsycINFO, and SCOPUS. There were no date or type of publication restrictions. Once duplicates were removed, this search produced 304 unique publications.

The following criteria were used to focus the selection, and was an adaptation of König et al.’s (2022) process: (1) publications had to be in a peer-reviewed publication, and (2) had to be published in English. This resulted in 137 publications excluded due to publication type, and 15 due to language. The third consideration was reviewing the titles, abstracts, and keywords to determine whether the use of ecolog* was within the scope of this project, meaning that ecolog* had to be used to describe a phenomenon, theory, framework, or approach (e.g., ecological approach, learning ecologies, ecocultural). If there was one instance of ecological being used in a comparable way, or if it could not be determined in the title, abstract, and keywords, the publication was retained at this step. If it was at all possible that the publication could address the project scope, it was left in for further review. Articles that were clearly only about ecology topics (e.g., ecological gardens, ecological diversity, ecological footprint, an ecology class) rather than ecological approaches or lenses were excluded at this step ($n = 77$). After considering criteria 1-3, 76 articles were carried forward to evaluate. Figure 1 shows the search and screening steps.

The full text of the 76 remaining articles were read and evaluated for three exclusion criteria (4-6): (4) the article was not about mathematics education. These articles were likely in the search given the use of math* and education, but addressed the two topics separately (e.g., math achievement scores in a social studies context) and were not about math education ($n = 11$). Exclusion criterion 5 was the article did not include an ecological lens. Thirteen of the publications that were left in due to ambiguity at criterion 3 were removed after a full-text review of how ecolog* was used. The same criterion was used for criterion 5 as criterion 3, namely publications that were only about ecology topics such as gardens and climate change were removed; there was no additional criterion added in this step compared to criterion 3. Exclusion criterion 6 was the article did not include a description of *ecological*. The publication had to provide a definition or explanation, of any length, to indicate what the author(s) meant by *ecological*. If this was not included, the publication was eliminated at this step since it could not reliably be determined what the author(s) intended ($n = 9$). At the conclusion of the search and selection process, 43 articles met all criteria and were retained for analysis.

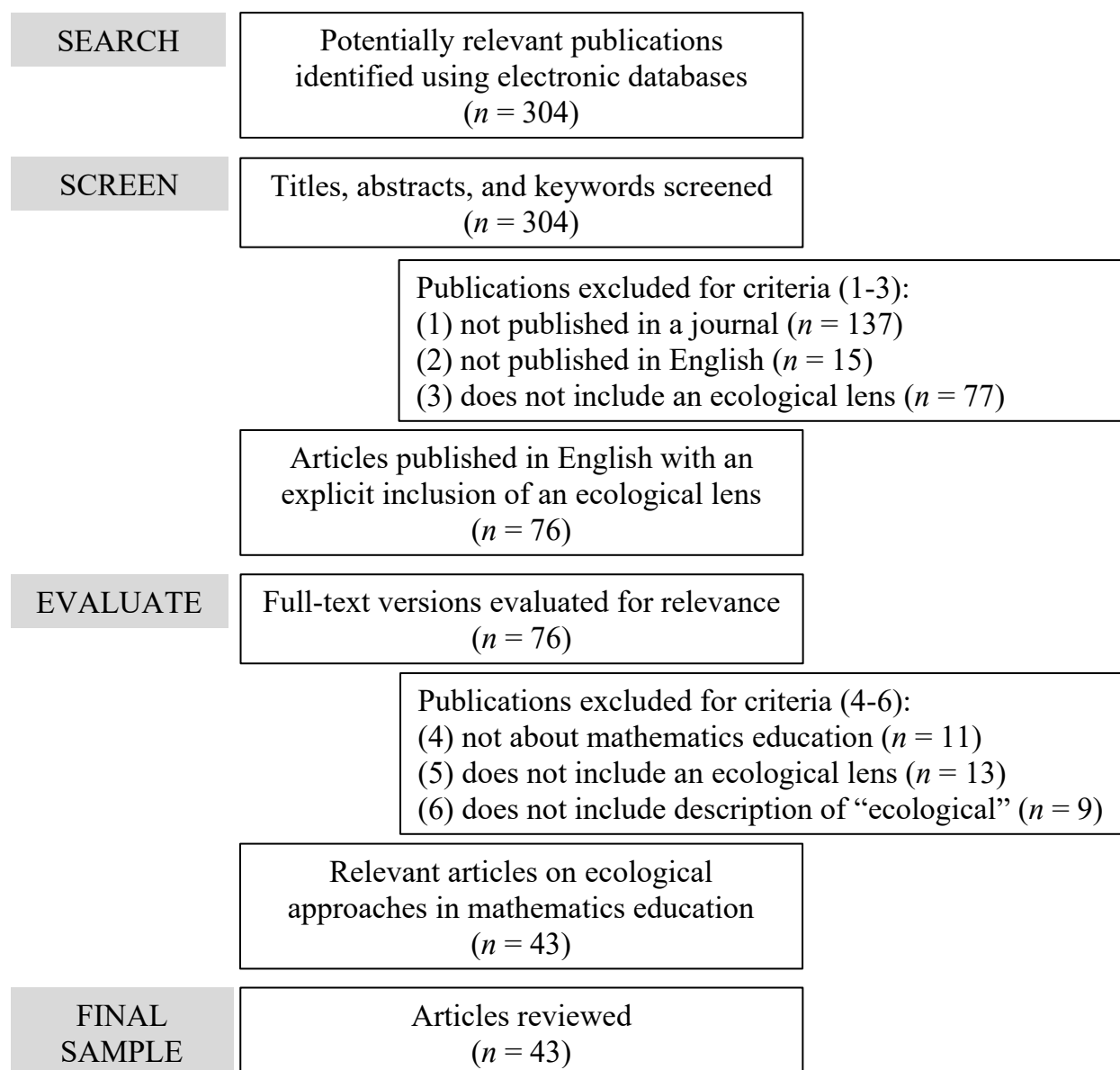


Figure 1: Search and selection process (adapted from König et al., 2022).

Descriptions and conceptualizations of “ecological approaches”

The 43 articles in mathematics education were highly variable in their use of *ecological* (and variations), and used 55 different phrases to incorporate ecological in their work. Most phrases ($n = 30$) were only used in one article, 17 phrases were used in two or three articles, and eight phrases were used in four to seven articles. This means there are many phrases seldom used, indicating researchers tend to use their own descriptive language associated with *ecological* and there is not consistency across mathematics teacher education research terminology. The three most-frequent phrases, appearing in seven articles each (16%) were classroom ecology or ecosystem; ecological systems framework, theory, or model; and ecological approach. Only one of these, referring to Bronfenbrenner’s (1979) ecological systems, had a consistent definition and use. “Ecological approach” is a claim authors made about their own work, but without a

consistent reference point or criterion. The third phrase, referring to a classroom as an ecosystem or ecology, was not consistently defined in terms of what comprised the classroom ecosystem.

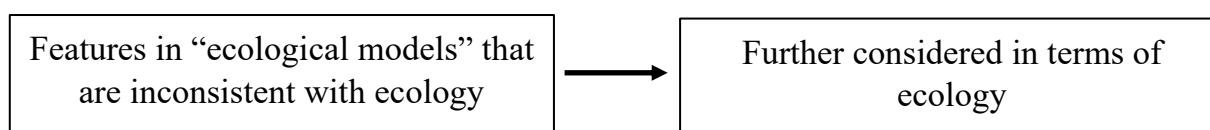
There was similar variability in conceptualizations, as the 43 publications used 32 different conceptualizations of *ecological*. The most frequently used was Bronfenbrenner's (1979) ecological systems theory or bioecological approach ($n = 10$). Second was Godino et al.'s (2007) onto-semiotic approach ($n = 9$), where one of six facets of didactic suitability is ecological suitability, defined as the "extent to which the teaching and learning process fits the educational project, the school and society, and take into account the conditioning factors of the setting in which it is developed" (p. 133). The third most-frequently used theory was Gibson's (1979) ecological approach to perception ($n = 5$). The remaining 29 conceptualizations were used in only one ($n = 26$) or two ($n = 3$) articles, again indicating an extensive variety and little consistency in mathematics education researchers' conceptualizations of projects that they claim have an ecological component.

Gaps with ecology

Significantly, ecological conceptualizations in mathematics teacher education rarely used ecology (again, meaning the discipline). Ecology was directly included, at any mention, and with the most generous reading possible, in seven of the articles (16%). In the vast majority of articles, *ecological* was used as an imprecise descriptor without any explanatory power. It was often used to convey relationship, most frequently as a synonym for *connected*, *related*, or *interdependent*, but most researchers were not actually drawing from the discipline of ecology in their work. This was evident in their definitions and explanations of their terminology, and was further evident by analysing the framework, approach, or model identified as *ecological*. Most of the time, their self-described "ecological model" or "ecological conceptualization" included elements that were incongruous with various principles in ecology. Figure 2 shows some of these features. The left side identifies features that were present in the "ecological models," but are inconsistent with ecology. The right side identifies how further considering ecology could provide shifts.

Insights from the articles that used ecology

Returning to the seven articles that used ecology, in four instances the authors considered the discipline of ecology through examples they included from "traditional environmental ecology" (Ehrenfeld, 2022, p. 491; Lupinacci & Happel-Parkins, 2017),



Identify elements in a system → Explore dynamic processes & relationships

Complicated set of pieces → Complex web of interactions

Fracture into discrete components → Maintain integrity of the work of teaching

Centre on an individual → Holistic understanding

Dualisms → Integration

Figure 2: “Ecological” conceptualizations found in the reviewed publications, further considered in terms of *ecology*.

their descriptions of the classroom as an ecology (Towers & Hunter, 2010), and their priority of “not taking the ecological (or ‘nature’) as fixed background for other concerns” (Coles, 2023, p. 21). The other three articles (Jazby et al., 2023; Scheiner, 2021; Videla et al., 2021) included ecology through Gibson’s (1979) ecological perception; it is unclear if ecology was wholistically considered beyond ecological psychology.

In each of these seven examples, *ecological* was not merely a descriptor but rather a foundational stance that was woven throughout the article. Towers and Hunter (2010) went as far as to describe their ecological perspective as “less a method of inquiry than an orientation to the world” (p. 28). Three core ideas that emerged from analysing the seven articles are as follows: (1) Shifting away from linear (Ehrenfeld, 2022; Scheiner, 2021) or mechanistic (Lupinacci & Happel-Parkins, 2017) models of cause and effect (Towers & Hunter, 2010), thinking (Lupinacci & Happel-Parkins, 2017), learning (Ehrenfeld, 2022), and knowing (Towers & Hunter, 2010) to instead focus on knowledge as participatory (Towers & Hunter, 2010) and thinking and meaning-making as relational (Lupinacci & Happel-Parkins, 2017). (2) De-centring the individual (Lupinacci & Happel-Parkins, 2017) as the “locus of concern” (Coles, 2023, p. 23), and instead focusing on *relationships* as the unit of learning (Coles, 2023), or “performer-environment *interaction*” (Jazby et al., 2023, p. 649; Scheiner, 2021) as the unit of analysis. (3) Weaving together the complex (Ehrenfeld, 2022) entanglement (Lupinacci & Happel-Parkins, 2017; Scheiner, 2021) of many interactions, including the social and physical (Lupinacci & Happel-Parkins, 2017; Videla et al., 2021), social and nature (Coles, 2023), nature and culture (Videla et al., 2021), language, relationships, and ideas (Towers & Hunter, 2010), that comprise, interact with, and elevate the significance of the environment (Jazby et al., 2023). The “relationship among the interconnectedness” of a “multitude of other living and non-living beings in an ecological system” (Lupinacci & Happel-Parkins, 2017, p. 57) is a primary focus.

DISCUSSION AND CONCLUSION

This study raises a broader set of questions, including when is it warranted for a researcher to claim their approach is ecological? Just because someone describes a classroom as an ecosystem, does that mean they were drawing from ecology, and is it warranted to describe their work as ecological? In this study, only one of seven articles that used the phrase “classroom ecology” or “classroom ecosystem” actually referred to ecology (Towers & Hunter, 2010). When ecological implications are used to develop one portion of a model, but set down to generate the remainder of the model, do we

consider that model ecological? What if a portion of the model is inconsistent with ecological principles? What is the saturation threshold that indicates a researcher has *used* versus *pointed to* (Helenius, 2024) a theory or discipline? My aim with this project is to present an invitation to conceptualize ecological approaches in mathematics teacher education research from the discipline of ecology, rather than using *ecological* as an imprecise descriptor, metaphor, or backdrop (Coles, 2023) that lacks specificity and conceptual richness, and avoids integrating ideas from nature itself.

Regarding teacher noticing, how do our understandings of noticing shift when we conceptualize noticing from the perspective of ecology? What happens when researchers *start with ecology directly*, instead of starting with ecological psychology and “going through” an intermediary like Gibson (1979) or Bronfenbrenner (1979)? There are likely multiple ways to conceptualize teacher noticing that are consistent with ecology. Seriously considering ecology and its integration into teacher noticing conceptualizations would mean using ecology throughout the entire model; this would entail, minimally, shifting away from linear models, de-centring an individual teacher, weaving together the complex entanglement of relationships and interactions, and *not* introducing elements that are incongruous with other principles of ecology. The unit of analysis or descriptive unit would shift to the dynamic processes and relationships among and between elements in the ecosystem. These would likely emerge in *episodes of interaction* rather than as an *outcome of individual cognition*.

Beyond teacher noticing, researchers across mathematics teacher education could consider their use of *ecological* and the consistency of their ideas and approaches with ecology, identify further opportunities to draw from ecology, and select when a different descriptor (e.g., *connected*, *related*) may better convey their thinking. Overall, this analysis revealed that ecological conceptualizations in mathematics teacher education rarely use ecology, and there are opportunities to further align conceptualizations labelled as *ecological* with ecology.

References

- Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.
- Buchanan, R. (2020). An ecological framework for supervision in teacher education. *Journal of Educational Supervision*, 3(1), 76–94.
- Coles, A. (2023). Towards a socio-ecological practice of mathematics teacher education. *AIEM – Avances de investigación en educación matemática*, 23, 19–35.
- Dindyal, J., Schack, E. O., Choy, B. H., & Sherin, M. G. (2021). Exploring the terrains of mathematics teacher noticing. *ZDM—Mathematics Education*, 53, 1–16.
- Ehrenfeld, N. (2022). Framing an ecological perspective on teacher professional development. *Educational Researcher*, 51(7), 489–495.
- Godino, J., Batanero, C., & Font, V. (2007.) The onto-semiotic approach to research in mathematics education. *ZDM—Mathematics Education*, 39, 127–135.

- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- Helenius, O. (2024, July 10). An extension of the theory of conceptual fields building on the concept of polysemy [Conference session]. 15th International Congress on Mathematics Education (ICME), Sydney, New South Wales, Australia.
- Jacobs, V. R., & Spangler, D. A. (2017). Core practices in K-12 mathematics teaching. In J. Cai (Ed.), *Compendium of research in mathematics education* (pp. 766–792). Reston: National Council of Teachers of Mathematics.
- Jacobs, V. R., Lamb, L. L. C., & Phillip, R. A. (2010). Professional noticing of children's mathematical thinking. *Journal for Research in Mathematics Education*, 41(2), 169–202.
- Jazby, D., Widjaja, W., Xu, L., & van Driel, J.H. (2023). Noticing student thinking under pressure in primary mathematics and science lessons. *International Journal of Science and Mathematics Education*, 21, 645–666.
- König, J., Santagata, R., Scheiner, T., Adleff, A. K., Yang, X., & Kaiser, G. (2022). Teacher noticing: A systemic literature review of conceptualizations, research designs, and findings on learning to notice. *Educational Research Review*, 36, 100453.
- Lupinacci, J., & Happel-Parkins, A. (2017). Ecocritically (Re)Considering STEM: Integrated ecological inquiry in teacher education. *Issues in Teacher Education*, 26(3), 52–64.
- Scheiner, T. (2021). Towards a more comprehensive model of teacher noticing. *ZDM—Mathematics Education*, 53, 85–94.
- Towers, J., & Hunter, K. (2010). An ecological reading of mathematical language in a Grade 3 classroom: A case of learning and teaching measurement estimation. *The Journal of Mathematical Behavior*, 29, 25–40.
- Videla, R., Aguayo, C., & Veloz, T. (2021). From STEM to STEAM: An enactive and ecological continuum. *Frontiers in Education*, 6, 709560.
- Weston, T. L., & Amador, J. M. (2023). Teacher noticing: A literature review of mathematics and science teacher noticing conceptualizations. *School Science and Mathematics*, 123(7), 293–308.

THE INFLUENCE OF MATHEMATICAL ACHIEVEMENT ON THE DEVELOPMENT OF INTUITIVE PROBABILISTIC REASONING IN LOWER SECONDARY SCHOOL STUDENTS

Leon Wiemann, Sascha Schroeder & Stefan Halverscheid

Georg-August-Universität Göttingen

Intuitive reasoning plays a crucial role in evaluating and comparing probabilities. This study investigates how mathematical achievement influences the development of probabilistic intuition among lower secondary school students in Germany. Using a questionnaire based on the work of Efraim Fischbein and a standardized mathematics test, Pearson correlations show that the relationship between mathematical achievement and probabilistic intuition increases with age. However, certain probabilistic biases, especially the Base-Rate-Fallacy and the Equiprobability Bias, persist as exceptions. These findings highlight the need for further research to understand and address these biases.

BACKGROUND

Students' Intuitive Misconceptions About Probability

Efraim Fischbein began to study students' probabilistic reasoning from an intuitive point of view. An intuition is defined as a subjectively obvious, holistic, inevitable and immediate cognition (Fischbein, 1975; 1987). In contrast to Piaget and Inhelder (1951), he suggested that children are able to develop primary intuitions about probability before they develop a system that allows them to perform mental operations. These intuitions are to be distinguished from secondary intuitions, which are formed in adolescence through a systematic process of instruction, such as probability courses in school (Fischbein, 1987). Primary intuitions are "not necessarily be overcome but rather replaced by secondary intuition" (Batanero et al., 2016, p. 17). Secondary intuitions are therefore characterized by formal knowledge becoming intuitive (Fischbein, 1987). To build such accurate intuitions, one of the main tasks of teaching probability is to confront students with their (primary) incorrect intuitions (Tsakiridou & Vavyla, 2015; Biehler & Engel, 2015).

In his later research, Fischbein focused on the development of intuitions in probabilistic reasoning. Fischbein treated these biases as "misconceptions" of probability (Fischbein & Schnarch, 1997), assuming that no transformation into an accurate secondary intuition has taken place. In recent literature, the term "preconception" is used instead (Batanero et al., 2016, p. 18). Fischbein and Schnarch (1997) studied the development of intuitive misconceptions in probabilistic reasoning in Israeli secondary school and university students between the ages of ten and twenty, using a questionnaire with well-known probabilistic preconceptions formulated by Amos Tversky, Daniel Kahneman

and others. The data show a complex situation in which some misconceptions decrease with age, while the frequency of others remains stable or even increases with age.

The relationship between the development of intuitive probabilistic reasoning and mathematical achievement

While other studies have shown that the structural teaching of probability can positively influence the development of secondary intuitions (Fischbein & Gazit, 1984; Rasfeld, 2004), and contemporary research has increasingly focused on importing frameworks and theories from other domains into mathematics education (Batanero et al., 2016, p. 18), less attention has been given to the role of individual differences in shaping probabilistic reasoning across age. Individual characteristics, such as mathematical achievement, prior knowledge and motivational factors, might influence intuitive probabilistic reasoning and its development.

This paper focuses on the influence of mathematical achievement on the development of intuitive probabilistic reasoning abilities. Looking at probabilistic reasoning from a conceptual knowledge point of view, several studies have shown that high mathematical achievement is associated with the students' ability in probabilistic reasoning. For example, Green (1982) found that 40 % of the variance in probabilistic reasoning ability among 2930 nine- to eighteen-year-old students in England can be explained by mathematical achievement. Similarly, Primi et al. (2017) reported a significant positive Pearson correlation ($r = .59, p < .001$) between mathematical ability and probabilistic reasoning in a sample of 317 undergraduate students in Italy.

However, these studies mainly emphasize formal approaches to reasoning, such as quantitative calculations of probabilities. He and Chen (2025) point out that such problems are “usually not designed to trick the participants' intuitions but to test their content and operational knowledge” (p. 147) by applying learned rules or formulas. Probabilistic intuitions may coexist with formal operations, potentially influence the application of learned procedures without being directly observable in arithmetically oriented problems (He & Chen, 2025). The aim of the study is to better understand how mathematical abilities and the development of probabilistic reasoning from an intuitive point of view are related.

RESEARCH QUESTION

The present study investigates the influence of mathematical achievement on the development of intuitive probabilistic reasoning. This connection is particularly important at the secondary school level, where fundamental ideas of probability, such as single and compound events, sample spaces, and probability distributions (Batanero et al., 2016), are systematically introduced. This is also implemented in the curricula in Germany (see Kultusministerkonferenz, 2022), where this study was conducted. At this stage, primary intuition in the area of probability needs to be replaced by secondary intuition in the sense of Fischbein (1975; 1987). The research question can be formulated as follows:

(Q1) How does students' mathematical achievement influence the development of their intuitive probabilistic reasoning over the age span in lower secondary school?

Based on the research question, it is hypothesized that a higher level of mathematical achievement is associated with a more developed level of intuitive probabilistic reasoning, as the formal and deductive knowledge acquired through (the teaching of) mathematics supports the formation of secondary intuitions.

RESEARCH DESIGN

The present study includes $N = 97$ lower secondary school students from two grammar schools (“Gymnasium”) in Germany. Of these, $n = 55$ students are attending the fifth grade (mean age: $M = 10.74$, $SD = 0.52$), while $n = 42$ students were visiting the seventh grade (mean age: $M = 12.66$, $SD = 0.48$).

No.	Title	Preconception
1	Base-Rate-Fallacy	Frequency of an event in the population is neglected (Kahneman & Tversky, 1973)
2+3	Insensitivity to Sample Size	Influence of sample size is neglected when comparing probabilities of events (Tversky & Kahneman, 1974; Sedlmeier & Gigerenzer, 1997)
4	Misconceptions of Chance	Probability of a series of outcomes is judged based on its representativeness (Tversky & Kahneman, 1974)
5	Negative Recency Effect	Probability of an event is overestimated due to recent observations (Tversky & Kahneman, 1974)
6	Bias of Imaginability	Frequency of an event is estimated based on the ease of generating examples (Tversky & Kahneman, 1974)
7+8	Conjunction Fallacy	Conjunction of two events is judged to be more probable than either single event due to its representativeness (Tversky & Kahneman, 1983)
9	Bayesian Reasoning	Conditional probability $P(A B)$ of two events A und B is confused with $P(B A)$ or $P(A \cap B)$ (Steib et al., 2025)
10	Time-Axis-Fallacy	Dependency is confused with causality (Fischbein & Schnarch, 1997)
11	Equiprobability Bias	Belief that probabilities of events tend to be equal (Lecoutre. 1992; He & Chen, 2025)
12	Outcome Approach	Characterization of probabilities as outcome of a single trial (Konold, 1995)

Table 1: Preconceptions used to measure students’ intuitive probabilistic reasoning

In terms of gender, $n = 46$ students identify as male, $n = 48$ identify as female, and $n = 3$ students did not report their gender. Students participated voluntarily and received € 10 (\approx \$ 10,30) for their participation.

Students' intuitive probabilistic reasoning is measured in both grades using a questionnaire originally presented in Schnarch (1995) and discussed in Fischbein and Schnarch (1997). The questionnaire, translated from Hebrew and English into German, consists of twelve items, all of which refer to well-known preconceptions of probability. In none of the items, students were required to compute probabilities. Instead, they were asked to compare the probability of certain events and to judge which event was most likely. Table 1 gives an overview of the questions and the corresponding preconceptions. The students' intuitive probabilistic reasoning scores will be evaluated based on the total number of correct answers.

To assess students' mathematical achievement, two versions of the "Deutscher Mathematiktest" (referred to as "DEMAT"; "German Test in Mathematics") were used. Fifth grade students were administered the DEMAT 5+ (Götz et al., 2013a), while those in the seventh grade received the DEMAT 6+ (Götz et al., 2013b). The administration of the tests was determined as specified in the test manuals. In Germany, both versions of the DEMAT are widely used to measure students' mathematical achievements (e. g., Preckel et al., 2017), offering high internal consistency (Cronbach's $\alpha = .89$ for DEMAT 5+, Götz et al., 2013a; Cronbach's $\alpha = .92$ for DEMAT 6+, Götz et al., 2013b). In addition, DEMAT 5+ and DEMAT 6+ make it possible to evaluate for students' performances in the subtests "Arithmetic" and "Applied Mathematics". "Arithmetic" is understood here as the use of basic arithmetic and algebraic operations in tasks without application. In contrast, "Applied Mathematics" requires the extraction, application and interpretation of mathematical information from texts, graphs and tables (Götz et al., 2013a; 2013b). In each subtest and the total score, students' achievement is normally distributed with $\mu = 50$ and $\sigma = 10$ (Götz et al., 2013a; 2013b).

RESULTS

To ensure the quality of students' responses to the intuitive probabilistic reasoning questionnaire, item difficulty, standard deviation and item discrimination are calculated for each item. The item difficulty ranges from .21 to .79. However, some items have low item discrimination, which refers to the ability of an item to distinguish between high and low scorers on the overall test. Primary item no. 11 ($r_{it} = .19$), that assesses Equiprobability Bias, and item no. 1 ($r_{it} = .26$), which evaluates the Base-Rate-Fallacy, are not strongly correlated with the sum of students' correct answers across all items. Previous research has shown that these preconceptions are among the most established and are particularly difficult to replace with accurate secondary intuitions (e.g., Kaplar et al., 2021; He & Chen, 2025). Therefore, it seems reasonable to exclude these items from the total score. Instead, they are treated independently below. In addition, item no. 2 (Insensitivity to Sample Size) is removed from the total

score due to its low item discrimination ($r_{it} = .29$). The students' intuitive probability score now includes the sum of the remaining nine correctly answered items. Grade five students achieved a mean score of $M = 3.82$ ($SD = 1.68$), while grade seven students answered $M = 4.74$ ($SD = 2.12$) items correct. The score demonstrates low internal consistency with Cronbach's $\alpha = .56$. A Shapiro-Wilk-Test reveal that students' intuitive probabilistic reasoning score for both grades significantly deviate from a normal distribution ($W = 0.958$, $p = .003$).

In both grades, students' mathematical achievement was assessed using the DEMAT. The total score, as well as the subtests "Arithmetic" and "Applied Mathematics", show acceptable to very good internal consistency with Cronbach's $\alpha \geq .63$. Besides, in grade five, students achieved a mean score of $M = 45.26$ ($SD = 9.64$), and in grade seven, the mean score was $M = 49.29$ ($SD = 10.24$). A Shapiro-Wilk test is performed for both grades, both subtests and the total score to verify whether the students' scores follow a normal distribution (Shapiro & Wilk, 1965). Except for DEMAT5+ Applied Mathematics, all test results support this assumption (DEMAT 5+ Arithmetic: $W = .970$, $p = .323$; DEMAT 5+ Applied Mathematics: $W = .941$, $p = .030$; Total DEMAT 5+ Score: $W = .950$, $p = .066$; DEMAT 6+ Arithmetic: $W = .954$, $p = .255$; DEMAT 6+ Applied Mathematics: $W = .945$, $p = .145$; Total DEMAT 6+ Score: $W = .976$, $p = .733$).

Table 2 presents Spearman's ρ for the students' intuitive probabilistic reasoning score, the Base-Rate-Fallacy and the Equiprobability Bias, separated by grade and the subtest of DEMAT. Since the intuitive probabilistic reasoning score is not normally distributed

Students' intuitive probability reasoning			
	Arithmetic	Applied Mathematics	Total DEMAT Score
Grade 5	.47**	.22	.41**
Grade 7	.48***	.61***	.65***
Base-Rate-Fallacy			
	Arithmetic	Applied Mathematics	Total DEMAT Score
Grade 5	-.13	-.04	-.12
Grade 7	.20	.17	.20
Equiprobability Bias			
	Arithmetic	Applied Mathematics	Total DEMAT Score
Grade 5	.16	-.20	.03
Grade 7	.13	.07	.14

** $p < .01$ *** $p < .001$

Table 2: Spearman's ρ by grade and (sub-)test of DEMAT

and the other two variables are binary, Spearman's ρ is used instead of Pearson's correlation coefficient. Spearman's ρ and associated p -values were computed using R (version 4.4.2) and the 'Hmisc' package.

DISCUSSION

In this study, a questionnaire by Fischbein and Schnarch (1997) was used to assess students' intuitive probability reasoning. While nine items are used to determine an intuitive probabilistic reasoning score, two items are focused isolated due to their low item discrimination: the Base-Rate-Fallacy and the Equiprobability Bias, which are known for their persistence as incorrect primary probabilistic intuitions. In addition, the mathematical achievement of fifth and seventh graders in Germany was evaluated using a standardized diagnostic test.

For fifth grade students, a moderately positive and significant correlation is observed between intuitive probabilistic reasoning and "Arithmetic". In contrast, for "Applied Mathematics", the correlation is weaker and not significant, which may be due to the non-normal distribution of the data for this subtest. Therefore, more data are needed to more reliably assess the connection between grade five students' intuitive probabilistic reasoning and their achievement in "Applied Mathematics". In grade seven, strong significant positive Spearman's ρ can be found for all subtests. The correlations are consistently stronger compared to grade five. While students' intuitive probabilistic reasoning score also increase from grade five to seven, this result also demonstrates that intuitive probabilistic reasoning aligns more closely with mathematical achievement across age.

In view of the research question of this the correlations indicates that a higher level of mathematical achievement is associated with accurate intuitive probabilistic reasoning. This association becomes stronger with age, especially in applied mathematics. Extracting and interpreting mathematical information therefore seems to foster the transformation of secondary intuitions in probability.

However, this does not appear to be the case for the Base-Rate-Fallacy and the Equiprobability Bias. No positive correlation coefficients were found for students' reasoning on these two items. Across both grades five and seven, both subtests, and the total DEMAT score, students' mathematical achievement does not seem to influence their intuition in these probabilistic problems as Spearman's ρ values for these items are all non-significant and close to zero, with the highest value not exceeding 0.20. On one hand, this limits the broader findings discussed earlier. On the other hand, it is consistent with previous research: the Base-Rate Fallacy, in which an population frequency of an event is ignored in favor of seemingly more representative information (Tversky & Kahneman, 1973), and the Equiprobability Bias, in which the probabilities of different events are assumed to be equal (Lecoutre, 1992), are both well-documented in various age groups and exceptionally difficult to overcome. For example, the Equiprobability Bias is the only intuition found to be stable across age in Fischbein

and Schnarch (1997). Further research is needed to explore, why these two biases in particular are not associated with mathematical achievement.

Empirical research is needed to determine whether emphasizing probabilistic concepts and problems in upper secondary school, particularly conditional probabilities and the law of large numbers (Batanero et al., 2016), enhances probabilistic intuition through improved mathematics performance. The questionnaire developed by Fischbein and Schnarch (1997) can also be administered to students in higher grades, such as grades nine and eleven, to further explore this potential relationship. In addition, other personal characteristics such as gender, interest, anxiety about mathematics or mathematical beliefs should be considered to better understand how preconceptions about intuitive probabilistic reasoning can be addressed and potentially overcome.

References

- Batanero, C., Chernoff, E. J., Engel, J., Lee, H. S., & Sánchez, E. (2016). *Research on Learning and Teaching Probability*. Springer Cham.
- Fischbein, E. (1975). *The intuitive sources of probabilistic thinking in children*. Reidel.
- Fischbein, E. (1987). *Intuition in science and mathematics: an educational approach*. Reidel.
- Fischbein, E., & Gazit, A. (1984). Does the Teaching of Probability Improve Probabilistic Intuitions? An Exploratory Research Study. *Educational Studies in Mathematics*, 15(1), 1-24.
- Fischbein, E., & Schnarch, D. (1997). The Evolution with Age of Probabilistic, Intuitively Based Misconceptions. *Journal for Research in Mathematics Education*, 28(1), 96-105.
- Götz, L., Lingel, K., & Schneider, W. (2013a). *Deutscher Mathematiktest für fünfte Klassen (DEMAT 5+)*. Hogrefe.
- Götz, L., Lingel, K., & Schneider, W. (2013b). *Deutscher Mathematiktest für sechste Klassen (DEMAT 6+)*. Hogrefe.
- He, S., & Chen, C. (2025). The Effects of Equiprobability Bias and Representativeness Heuristics on the Performance in Probability Comparison and Calculation Tasks Among Middle School Students in China. *International Journal of Science and Mathematics Education*, 23(1), 143-168.
- Kahneman, D., & Tversky, A. (1973). On the Psychology of Prediction. *Psychological Review*, 80(4), 237-251.
- Kaplar, M., Lužanin, Z., & Verbić, S. (2021). Evidence of probability misconception in engineering students—why even an inaccurate explanation is better than no explanation. *International Journal of STEM Education*, 8(18), 1–15.
- Konold, C. (1995). Issues in Assessing Conceptual Understanding in Probability and Statistics. *Journal of Statistics Education*, 3(1), 1-9.
- Kultusministerkonferenz (2022). *Bildungsstandards für das Fach Mathematik. Erster Schulabschluss (ESA) und Mittlerer Schulabschluss (MSA)*. Kultusministerkonferenz.

- Lecoutre, M. P. (1992). Cognitive models and problem spaces in “purely random” situations. *Educational Studies in Mathematics*, 23(6), 557–568.
- Piaget, J., & Inhelder, B. (1951). *La genèse de l'idée de hasard chez l'enfant*. Presses Universitaires de France.
- Preckel, F., Schmidt, I., Stumpf, E., Motschenbacher, M., Vogl, K., & Schneider, W. (2017). A Test of the Reciprocal-Effects Model of Academic Achievement and Academic Self-Concept in Regular Classes and Special Classes for the Gifted. *Gifted Child Quarterly*, 61(2), 103-116.
- Primi, C., Morsanyi, K., Donati, M. A., Galli, S., & Chiesi, F. (2017). Measuring Probabilistic Reasoning: The Construction of a New Scale Applying Item Response Theory. *Journal of Behavioral Decision Making*, 30(4), 933-950.
- Schnarch, D. (1995). *The Evolution with Age of Probabilistic, Intuitively Based Misconceptions*. Tel-Aviv University.
- Sedlmeier, P., & Gigerenzer, G. (1997). Intuitions About Sample Size: The Empirical Law of Large Numbers. *Journal of Behavioral Decision Making*, 10(1), 33-51.
- Shapiro, S. S., & Wilk, M. B. (1965). An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, 52(3-4), 591-611.
- Tversky, A., & Kahneman, D. (1974). Judgement under Uncertainty: Heuristics and Biases. *Science*, 185(4157), 1124-1131.
- Steib, N., Büchter, T., Eichler, A., Binder, K., Krauss, S., Böcherer-Lindner, K., Vogel, M., & Hilbert, S. (2025). How to teach Bayesian reasoning: An empirical study comparing four different probability training courses. *Learning and Instruction*, 95, 102032, 1-18.
- Tsakiridou, H., & Vavyla, E. (2015). Probability concepts in primary school. *American Journal of Educational Research*, 3(4), 535-540.
- Tversky, A., & Kahneman, D. (1982). Judgement under uncertainty: Heuristics and biases. In D. Kahneman, P. Slovik & A. Tversky (Hrsg.), *Judgement under uncertainty: Heuristics and Biases* (pp. 3-20). Cambridge University Press.
- Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgement. *Psychological Review*, 90(4), 293-315.

LIKES, VIEWS, COMMENTS: HOW IS VIEWER ENGAGEMENT RELATED TO HIGH- AND LOW-QUALITY EXPLANATORY VIDEOS?

Laura Wölck, David Bednorz, Aiso Heinze

IPN – Leibniz Institute for Science and Mathematics Education, Kiel (Germany)

Mathematical explanatory videos on platforms like YouTube are widely used by learners but often criticized by experts for their quality. Research on the relationship between viewer engagement and the educational quality of videos is limited and shows inconsistent results. Using existing quality data from 44 YouTube videos on the derivative concept, we examined differences in engagement metrics between high- and low-quality videos. Results reveal a nuanced relationship between video quality and viewer engagement. While no strong differences were found, effect sizes suggest meaningful variations, with low-quality videos showing higher appreciation and interactive engagement of viewers, but high-quality videos demonstrating higher long-term consumption rates.

INTRODUCTION

In recent years, the importance of explanatory videos has increased significantly, both as a learning medium in schools and as a self-chosen supplement for learning outside the classroom. Especially mathematical explanatory videos have gained high popularity among learners. Platforms like YouTube offer a diverse range of explanatory videos, making it possible to find relevant content for almost any topic. However, YouTube allows anyone to publish explanatory videos, causing an increasing concern about the content quality on YouTube (Shoufan & Mohamed, 2022). Studies reveal that mathematical YouTube videos often depict mathematics negatively, focus on procedures over concepts, lack context, and frequently contain errors and inconsistent terminology (Bersch et al., 2020; Wölck et al., under review).

Despite these issues, not all explanatory videos are of poor quality, raising questions about how to identify high-quality content and understand student preferences in terms of engagement. YouTube provides surface-level metrics that can be used to measure viewer engagement and video popularity, including the number of likes, comments, and views. These metrics not only reflect how well a video resonates with its audience but also serve as indicators of its overall reach and reception (Saurabh & Gautam, 2019). Nevertheless, a key question remains unanswered: Is viewer engagement on YouTube related to the quality of mathematical explanatory videos as assessed by experts? This question is particularly relevant because students may prioritize different quality aspects than mathematics education experts or fail to recognize certain errors, leading them to favor videos that experts might critique. Previous research on the relationship between surface metrics, such as viewer engagement, and video quality

has produced inconsistent results (Bitzenbauer et al., 2024; Kulgemeyer & Peters, 2016; Wölck et al., under review). This study aims to explore this question further and provide new insights into this area of research.

THEORETICAL BACKGROUND

Quality of mathematical explanatory videos

Studies analyzing selected mathematical videos on YouTube have revealed several concerning trends. These videos often portray mathematics negatively, emphasize procedural over conceptual knowledge, present content without meaningful context, and frequently contain factual errors as well as inconsistent use of technical language (Bersch et al., 2020). In this respect, Wölck et al. (under review) analyzed over 1,000 YouTube videos on derivatives - a fundamental calculus concept essential for advanced mathematics and various applications - finding most videos emphasizing procedural over conceptual knowledge. However, videos teaching conceptual knowledge are particularly important, as acquiring such knowledge poses a challenge for many students. Students often struggle with understanding derivatives, relying on memorization over conceptual understanding (Thompson & Harel, 2021).

To investigate the quality of explanatory videos in mathematics that convey conceptual knowledge, a reliable and valid quality score is required. Therefore, Wölck et al. (under review) developed an elaborated measurement instrument for the educational quality of explanatory videos, which addresses general educational, overarching mathematical educational, and topic-specific categories. The study evaluated 44 YouTube videos on the concept of derivative of the most popular mathematical YouTube channels in German. The quality assessment of these videos revealed heterogeneous results, with the occurrence of errors in several videos and many videos deemed "poor" concerning the topic-specific quality criteria on teaching the derivative.

Video quality and viewers engagement

Recent research, particularly in science education, has explored viewers engagement with different metrics as indicators and the relationship with subject-specific video quality for explanatory videos. One question is whether, especially for teachers and students, the time-consuming nature of quality assessment using detailed criteria can be replaced by quicker and easier-to-collect indicators. According to Reinhold et al. (2024), students are to a certain extent able to perceive quality criteria in mathematics videos under controlled conditions of video comparison. The question arises to what extent this also works when students choose videos freely on platforms like YouTube.

The current state of research on this topic is limited. There are two studies investigating explanatory videos in physics-related topics, yielding heterogeneous results. Kulgemeyer and Peters (2016) analyzed explanatory videos for classical mechanics topics and found that only the number of content-related comments was correlated with explaining quality but not the number of views or likes. Bitzenbauer et al. (2024) examined explanatory videos on quantum topics and extended previous results by

identifying a correlation between the number of likes and explaining quality. Additionally, their research suggested that the total number of interactions, such as the sum of likes and dislikes, might serve as an indicator of high explaining quality. For mathematical explanatory videos, Wölck et al. (under review) did not find correlations between viewer engagement (number of likes, views, or comments) and quality for the 44 videos on the concept of derivative.

PRESENT STUDY

The previously mentioned studies revealed mixed results, indicating that the relationship between viewer engagement and video quality may depend on factors such as subject matter or video characteristics. In addition, correlation analyses may not be sufficiently differentiated to capture the complex relationships between the considered variables. A more nuanced approach, comparing distinct quality groups of videos (e.g., the top and bottom terciles) in terms of viewer engagement, could reveal patterns that correlational studies might miss. Furthermore, examining various aspects of video quality separately, such as media design, cognitive activation, constructive support, and topic-specific criteria, rather than relying solely on an overall quality score, may provide more detailed insights. Accordingly, our study aims to employ a multivariate approach to investigate potential differences in surface metrics between high- and low-quality mathematical explanatory videos across various quality dimensions by addressing the following research question:

RQ: How engaged are viewers of mathematical explanatory videos in high-quality or low-quality videos, considering both overall quality and specific quality dimensions (general criteria, overarching mathematics educational criteria, topic-specific criteria)?

METHOD

We used the existing data of the 44 explanatory videos on the concept of derivative and the associated quality scores from the study of Wölck et al. (under review). All videos were retrieved from YouTube between December 2023 and March 2024. Video selection criteria focused on titles related to the introduction of the derivative concept and the most prominent channels. Priority was given to videos that aimed to cover the entire concept, ideally encompassing the three sub-concepts difference quotient, differential quotient, and the derivative function. In cases where multiple suitable videos were available in a channel, the one with the highest view count was chosen.

The assessment of video quality was based on a theory-based criteria catalogue. The development of the criteria catalogue involved a review of already existing criteria in literature, incorporating insights from research on teaching quality and educational psychology, and adapting these principles to the context of explanatory videos. Additionally, the researchers drew upon studies on learning with explanatory videos, including the work of Mayer (2014), and specifically examined mathematics educational research on learning the concept of derivative. The resulting criteria catalogue encompasses the main categories: media design, cognitive activation,

constructive support, mathematics-specific criteria, and topic-specific criteria related to the concept of derivative. In total, the instrument comprises 62 criteria. The criteria catalogue has demonstrated objectivity, reliability, and validity in its application, ensuring its suitability for assessing the quality of mathematical explanatory videos (see Wölck et al., under review for more details on video selection and the quality rating). For the present study, the first three categories were combined to the group of *general quality criteria* and the last two categories to *mathematics-/topic-specific criteria*. The latter includes, for example, the use of appropriate representations and examples, correctness, and precise mathematical language.

Measuring viewer engagement

For this study, we operationalize viewer engagement through multiple dimensions to capture different aspects of user interaction with mathematical explanatory videos. We distinguish between two primary forms of engagement: video consumption and video appreciation. Video consumption is quantified using two metrics: the number of views and the number of comments. While views indicate how often a video was accessed, they do not necessarily reflect whether the video was watched in its entirety, as even brief interactions count as views. Including comments as an additional metric provides further insight into viewer behavior, as leaving a comment suggests a deeper level of engagement with the content, potentially indicating that the video was watched for a longer duration. However, the number of comments reflects consumption rather than appreciation, as comments can be both positive and negative. Video appreciation is measured by the number of likes a video receives, representing a more explicit form of positive engagement. By considering these distinct aspects, we aim to provide a comprehensive assessment of how viewers engage with mathematical explanatory videos of varying quality. This multifaceted approach allows for a nuanced analysis of engagement patterns across different quality dimensions.

To examine the relationship between assessed quality and engagement, various surface-level metrics for each video were collected in August 2024 from the YouTube platform. These metrics included the number of likes, views, comments, and the publication date. Using the publication date, the duration of each video's online presence in years was calculated.

Data analysis

To address the disparities in exposure duration and audience reach among the videos, we implemented a standardization process for the surface metrics that reflect viewer engagement. We calculated a standardized view rate by dividing the total views by the number of years the video had been online, enabling a fair comparison of video consumption irrespective of their publication dates. Additionally, we normalized likes and comments by expressing them as ratios to the total view count, facilitating a more equitable assessment of video appreciation and interactive engagement across videos with varying visibility levels. These standardization methods allow for more robust

comparisons of engagement patterns between videos by mitigating the influence of online duration and overall popularity on engagement metrics.

To analyze the relationship between video quality and viewer engagement, we conducted a multivariate analysis of variance (MANOVA) approach. We divided the videos into two groups: the bottom 33% (Group 1: $n = 15$ low-quality videos) and the top 33% (Group 2: $n = 15$ high-quality videos) based on their quality scores. This enables us to compare extremes, potentially revealing differences that might be obscured when examining the entire range of quality scores. The MANOVA was conducted to determine if these two groups differ significantly in terms of their standardized surface metrics (likes, views, and comments ratios). The analysis was performed not only for the overall quality score but also for two specific quality categories: general quality criteria and mathematics-/topic-specific criteria. This approach accounts for the possibility that students may place particular emphasis on specific aspects of video quality (mathematics-related or -unrelated aspects).

RESULTS

To investigate the relationship between video quality and viewer engagement, a series of MANOVAs were conducted. The independent variable was the quality group, categorized based on total quality score, general quality score and mathematics-/topic-specific quality score. The dependent variables included likes-to-views ratio, comments-to-views ratio, and views-to-duration ratio. The MANOVAs revealed no significant multivariate effects across all scoring categories. Specifically, the total score showed Pillai's Trace = .13 ($F(3,26) = 1.29, p = .30$) and the general score Pillai's Trace = .02 ($F(3,28) = 0.22, p = .88$). The mathematics-/topic-specific score yielded the highest Pillai's Trace of .16 ($F(3,27) = 1.73, p = 0.19$), though still not statistically significant. Follow-up univariate ANOVAs demonstrated no significant differences between groups for any dependent variables, however, the effect sizes, estimated by partial eta-squared (η^2), revealed noteworthy magnitudes despite the lack of statistical significance. According to Cohen's (1988) guidelines for interpreting η^2 , the total quality score ($\eta^2 = .13$) indicated a medium to large effect, while the mathematics-/topic-specific score ($\eta^2 = .16$) suggested a large effect. The general score ($\eta^2 = .02$) showed a small effect. These effect sizes suggest potentially meaningful differences between high- and low-quality videos in terms of viewer engagement, particularly for the total and mathematics/topic-specific quality scores. However, the absence of statistical significance, possibly due to the small sample size, necessitates cautious interpretation of these findings. If we look at the descriptive values in Table 1, they reveal interesting patterns of viewer engagement across different quality dimensions of mathematical explanatory videos.

Regarding video appreciation, as measured by the likes-to-views ratio, low-quality videos consistently showed higher engagement across all quality categories. This trend was mirrored in the comments-to-views ratio, an indicator of interactive engagement, where low-quality videos again demonstrated higher levels of engagement. However,

when examining video consumption through the views-to-duration ratio, a different pattern emerged. For both the overall quality and mathematics-/topic-specific criteria categories, high-quality videos exhibited substantially higher viewing rates over time. Interestingly, this trend was reversed for the general criteria category, where low-quality videos showed a higher views-to-duration ratio. These findings suggest a complex relationship between video quality and viewer engagement. While low-quality videos appear to elicit more explicit forms of engagement (likes and comments), high-quality videos tend to accumulate more views over time, particularly when considering overall quality and mathematics-/topic-specific quality. This discrepancy suggests the multifaceted nature of engagement and that different aspects of video quality may influence various forms of viewer interaction differently.

Quality Category	Engagement Metric	Low-Quality Videos <i>M (SD)</i>	High-Quality Videos <i>M (SD)</i>
Total quality criteria	Likes	0.024 (0.03)	0.017 (0.01)
	Comments	0.003 (0.005)	0.001 (0.001)
	Views	5,679 (13,604)	13,845 (23,941)
General quality criteria	Likes	0.024 (0.02)	0.020 (0.01)
	Comments	0.003 (0.004)	0.002 (0.003)
	Views	8,199 (16,209)	5,746 (7,211)
Mathematics-/topic-specific quality criteria	Likes	0.024 (0.03)	0.022 (0.01)
	Comments	0.003 (0.004)	0.001 (0.001)
	Views	6,375 (13,503)	13,282 (24,166)

Table 1: Descriptive Statistics for Engagement Metrics by Video Quality and Quality Category Note. *M* = Mean; *SD* = Standard Deviation. The engagement metrics (likes, comments, views) represent the standardized metrics.

DISCUSSION

The research question regarding viewer engagement with high-quality versus low-quality mathematical explanatory videos does not yield conclusive results. Statistical analyses through MANOVAs revealed no significant differences in engagement metrics between high- and low-quality videos across all quality dimensions. This finding aligns with previous research by Wölck et al. (under review) on mathematical videos and Kulgemeyer and Peters (2016) on physics videos. However, the sample size is small and effect sizes suggest potentially meaningful differences, particularly for total quality ($\eta^2 = .13$) and mathematics-/topic-specific criteria ($\eta^2 = .16$).

Descriptive statistics allow detailed information on engagement patterns:

1. Video appreciation: Low-quality videos consistently received more likes across all quality categories.

2. Video consumption: High-quality videos demonstrated higher views-to-duration ratios for overall and mathematics-/topic-specific criteria, suggesting greater viewership. Low-quality videos showed higher comments-to-views ratios, indicating more discussion. While comments can be both positive and negative, their presence at least suggests that viewers engaged more deeply with the content and likely watched the videos for a longer duration.

These findings indicate that viewer engagement might differ between high- and low-quality videos, but not uniformly across all engagement metrics. Low-quality videos appear to elicit more explicit engagement (likes and comments), while high-quality videos tend to accumulate more views over time. If this discrepancy can be replicated, it would point to a multifaceted nature of engagement in educational videos and suggest that different quality aspects may influence various forms of viewer interaction differently. The lack of statistical significance, possibly due to limited sample size, necessitates cautious interpretation. Further research with larger samples is needed to confirm these engagement patterns and their relationship to video quality dimensions.

One must also keep in mind, that the complex relationship between video quality and viewer engagement on YouTube is influenced by several factors beyond the content itself. The YouTube algorithm significantly affects video visibility and engagement patterns. High-quality explanatory videos may not receive more views or interactions if they are not effectively promoted by the algorithm, which prioritizes factors like watch time and viewer engagement metrics. User behavior also plays a crucial role. Viewers may engage with videos based on factors unrelated to educational quality, such as entertainment value, video length, channel popularity, or personal preferences. Additionally, the criteria used to evaluate video quality may not align with what drives viewer engagement on the platform. While educational quality focuses on accuracy and depth, the algorithm often rewards videos that achieve high watch time and interaction rates. This interplay between content quality, user behavior, and platform algorithms creates a complex landscape for viewer engagement.

Our study comes with several limitations. We considered only 44 videos, all of them addressed the topic derivative and each belong to a video channel in German providing other mathematical videos. Although the quality was recorded using an elaborate system with 62 criteria, relevant quality aspects may have been ignored. Finally, viewer engagement was only based on the number of likes, views, and comments.

Kulgemeyer and Peters (2016) suggest a promising approach of examining the nature of comments. This method could be applied to mathematical videos in future studies. However, given the quantitative focus of our current research, a qualitative analysis of comments was not conducted. Future investigations could incorporate this qualitative aspect to provide deeper insights into viewer engagement and perceptions of mathematical educational videos. Furthermore, it would be valuable to investigate the actual learning outcomes associated with these videos. Given that even lower-quality videos receive likes and views, it is crucial to examine whether students can still

acquire knowledge from such content. This approach would provide a more comprehensive understanding of the relationship between video quality, viewer engagement, and educational effectiveness.

References

- Bersch, S., Merkel, A., Oldenburg, R. & Weckerle, M. (2020). Erklärvideos: Chancen und Risiken [Explanatory videos: Opportunities and risks]. *Mitteilungen der GDM*, 109, 58–63. <https://ojs.didaktik-der-mathematik.de/index.php/mgdm/article/view/966>
- Bitzenbauer, P., Höfler, S., Veith, J. M., Winkler, B., Zenger, T. & Kulgemeyer, C. (2024). Exploring the Relationship Between Surface Features and Explaining Quality of YouTube Explanatory Videos. *International Journal of Science and Mathematics Education*, 22, 25–48. <https://doi.org/10.1007/s10763-022-10351-w>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Hillsdale, N.J: L. Erlbaum Associates.
- Kulgemeyer, C. & Peters, C. H. (2016). Exploring the explaining quality of physics online explanatory videos. *European Journal of Physics*, 37(6), 65705. <https://doi.org/10.1088/0143-0807/37/6/065705>
- Mayer, R. E. (2014). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 31–48). Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.005>
- Reinhold, F., Flubacher, R., & Mindl, V. (2024). Do quality criteria of instructional explanatory videos in algebra influence student choice? A feasibility study using conjoint analysis to assess implicit reasoning in educational contexts. *Mathematical Thinking and Learning*, 1–12. <https://doi.org/10.1080/10986065.2024.2377423>
- Saurabh, S. & Gautam, S. (2019). Modelling and statistical analysis of YouTube's educational videos: A channel Owner's perspective. *Computers & Education*, 128, 145–158. <https://doi.org/10.1016/j.compedu.2018.09.003>
- Shoufan, A. & Mohamed, F. (2022). YouTube and Education: A Scoping Review. *IEEE Access*, 10, 125576–125599. <https://doi.org/10.1109/access.2022.3225419>
- Thompson, P. W. & Harel, G. (2021). Ideas foundational to calculus learning and their links to students' difficulties. *ZDM – Mathematics Education*, 53(3), 507–519. <https://doi.org/10.1007/s11858-021-01270-1>
- Wölck, L., Bednorz, D. & Heinze, A. (under review). Beyond Views and Likes: Assessing the Didactic Quality of Mathematical Explanatory Videos Based on General and Topic-Specific Criteria.

STUDENTS' SELF-ASSESSMENT IN THE FIELD OF FRACTIONS: HOW ACCURATE ARE THEY?

Alexander Wolff¹, David Bednorz², and Barbara Schmidt-Thieme¹

¹University of Hildesheim, Germany; ²IPN Leibniz Institute for Science and Mathematics Education, Germany

Self-assessment is a central component of self-regulated learning in mathematics. For this, students must be able to assess themselves accurately, ensuring that their self-assessment aligns with their actual performance. To investigate the accuracy of students' self-assessment in the field of fractions, we conducted a quantitative empirical study with 214 students. The results show that students can assess themselves overall accurately. However, when looking at the different underlying skills, it showed that the accuracy varies between the skills requiring the mastery of procedures and those requiring the mastery of application. While students can rather assess themselves accurately regarding procedures of fractions, they tend to be less accurate in their self-assessment of the skill of applying fractions.

INTRODUCTION

Self-assessment can play a central role in learning mathematics, especially when studying for tests and exams or during periods of independent learning at school or at home. As an aspect of self-regulated learning, self-assessment is crucial for addressing the increasing challenges of heterogeneous conditions in teaching. It supports the individualization of learning and promotes lifelong learning (Hoyle & Dent, 2018; Zimmerman, 2002). Consequently, self-assessment is also required to be included in curricular requirements (NCTM, 2000). It aims to assess students' priors as basis for further learning processes. Whereby, self-assessment can only be effective, if students can assess their priors accurately. This means that the assessment of the priors corresponds to the requirements of the learning activities. If there is a mismatch between one's priors and the requirements of the learning activities, students could be over- or under-challenged, which could have an adverse effect on further learning processes (Raaijmakers et al., 2019; Brown & Harris, 2013).

To investigate the accuracy of students' self-assessment, we chose the topic of fractions. Many students have difficulties with this concept, which is fundamental for advanced topics in secondary school mathematics such as probability calculation or ray theorems. Due to its importance, the concept of fractions is repeatedly practiced throughout the school years.

For our investigation, we developed a quantitative test that consists of two parts. In the first part, we asked students how sure they are in mastering different skills related to the concept of fractions. In the second part, we asked students to complete them. In order to find out, how accurately the students can assess themselves, we examined the

connection between the two respective constructs that were asked. Our objectives to find out are: (1) How accurately can students assess themselves overall in the field of fractions? And (2) how accurately can students assess themselves in the different skills in the field of fractions?

SELF-ASSESSMENT AS A PART OF SELF-REGULATED LEARNING

Self-assessment is a central component in the development of self-regulated learning. Self-regulated learning can be defined as a goal in a constructivist mathematics classroom (see De Corte et al., 2000; Stillman & Mevarech, 2010; NCTM, 2000) and can contribute to a deeper, more meaningful learning process and thus to higher learning outcomes (Valle & Andrade 2014, Brown & Harris 2014). It refers to an active and independent learning process, in which students plan, monitor, adjust and evaluate themselves to effectively achieve their learning goals (Low & Jin, 2012; Zimmerman, 1998). Therefore, self-assessment takes places at the beginning of the process of self-regulated learning (Andrade & Valcheva, 2009; Schunk, 1996). In self-assessment students must determine what they already know to take appropriate measures to control the further learning process. Instead of self-assessment, terms such as self-estimation, self-evaluation, self-grading and others are sometimes used interchangeably to refer to the process of reflecting on and evaluating students' own performance for learning (Brown & Harris, 2013; Valle & Andrade, 2014). Following Andrade and Valcheva (2010, p. 13), we understand self-assessment as a "a process of formative assessment during which students reflect on the quality of their work, judge the degree to which it reflects explicitly stated goals or criteria, and revise accordingly". We differentiate self-assessment from self-concept, although both have similarities in their definition. In contrast to self-concept, we define self-assessment as task-related and therefore based on specific criteria (Andrade, 2010; Brown & Harris, 2013). We understand accuracy as the correspondence between self-assessment and the actual performance in a task.

SELF-ASSESSMENT IN THE FIELD OF FRACTIONS

Fractions as a fundamental mathematical concept form the basis for other topics and math in everyday life. Furthermore, knowledge of fractions is an important prerequisite for other learning subjects (Behr et al., 1983; Ohlsson, 1988; Obersteiner et al., 2019). Accordingly, fractions occur as a learning objective across the school years (Obersteiner et al., 2019). In addition to their general importance, fractions are traditionally part of the school curriculum that students find comparatively difficult (Lamon, 2007; Obersteiner et al., 2019a; 2019b; Lortie-Forgues et al., 2015). Moreover, fractions are a complex topic that encompasses a variety of different skills. These skills can be categorized into: adding fractions, subtracting fractions, multiplying fractions, dividing fractions, expanding and reducing fractions. They can also be distinguished according to whether they concern the mastering of procedures within the formal-symbolic notation, or whether they concern the application of the mathematical concept to the real world in the form of word problems (Behr et al.,

1983). Mastering procedures and applications reflect typically two aspects of learning mathematics (NCTM, 2000).

In previous cross-disciplinary studies, Brown and Harris (2013) as well as Andrade (2019) report in their meta-analysis an accuracy of $r=0.2-0.8$ for students' self-assessment. Among these, there are few studies that report accuracy higher than 0.6 (Brown & Harris, 2013; Andrade, 2019). Regarding Bradshaw (2001) and Hewitt (2005), it can be noticed that for more complex tasks, the accuracy of self-assessment is lower. They attribute this to the fact that the complexity is caused by the consideration of several relevant criteria. Andrade (2019) also argues that many studies examine students' self-assessments by relating them to teachers' assessments rather than to specific criterion-based tasks. In mathematics education, there are only few studies that have examined self-assessment such as Stalling and Tascione (1996), Ross et al. (2002) and Ramdass and Zimmerman (2008). However, there is no study to date that has examined self-assessment regarding fractions, although the concept is as central as mentioned. Therefore, we address the following questions: (1) How accurately can students assess themselves overall in the field of fractions? (2) How accurately can students assess themselves for the different skills in the field of fractions? With the first question, we aim to investigate how students can assess themselves for fractions as a whole. For this purpose, we combine and refer to the forehead mentioned skills together. With the second question, we want to examine the accuracy of self-assessment for the different skills of fractions. This includes the investigation of the accuracy of students' self-assessment regarding tasks with procedures and application. We expect students' self-assessment to be more accurate for procedural tasks than for application tasks, as application tasks are more complex (Pongsakdi et al., 2019; Prediger & Wessel, 2013). To answer the questions, accuracy is measured with a task-based self-assessment related to actual task performance.

METHOD

We collected data from two secondary schools that included a total of 214 students. The mean grade in mathematics was $M=3.04$ ($SD=0.98$), which indicates that students' performance was moderately satisfactory overall (1=very good; 6=insufficient). These students are distributed across school years 7-10 as follows: school year 7: 78; school year 8: 59; school year 9: 46; and school year 10: 22.

A quantitative performance test was developed and administered to answer the questions. The test consisted of two parts. In the first part the students were presented with tasks relating to six different skills and were asked to rate how sure they felt about themselves in this regard. For this purpose, four response options were provided on the response scale, ranging from: not at all sure, rather not sure, somewhat sure and completely sure. This scale level was selected to appropriately reflect the expected response behaviour. In the second part of the test, the students were asked to complete the tasks on the six different skills. The data was coded using a dichotomous scale of 0 and 1. The second part of the test could only be started when the first part had been completed.

Congruent scales were created, i.e. scales that fit each other based on the skill and thus the items. The tasks in the second part were not identical to the tasks in the first part but had the same structure. For example, instead of $\frac{1}{4} + \frac{1}{4}$, $\frac{1}{3} + \frac{1}{3}$ was chosen. The total scale of self-assessment and of performance items consists respectively of 40 items each, with Cronbach's alpha values of 0.98 and 0.93 respectively. The scales for self-assessment and performance items for skills two to six consisted of 7 items each, while the scale for skill seven consisted of 5 items. The Cronbach's alpha values for the self-assessment constructs regarding the different skills are between 0.92 and 0.95. The Cronbach's alpha values for the performance constructs of the various skills are between 0.75 and 0.91. Based on the amount, the diversity and the content representativeness, we interpret our Cronbach alpha values between good to very good (Streiner, 2003). For the research questions, we used the statistical methods of intraclass correlation in the field of classical test theory. To do this, we created aggregated mean scales from the items, which represent the constructs. This allowed us to calculate intraclass correlation values between the constructs. Doing so, we z-standardized the constructs beforehand to bring them to the same scale level. To analyze the data, we considered self-assessment and actual performance as raters in the formula. Therefore, we calculated the average score intraclass correlation coefficient (ICC). The model was based on a two-way approach (two-way analysis of variance) and measures the agreement of values (type: agreement).

In previous studies (see Brown & Harris, 2013; Andrade, 2019), the correlation was used as metric to assess the accuracy of self-assessment, which we found inadequate, because correlation is a measure of association rather than a measure of agreement. Regarding the interpretation of intraclass correlation values Cicchetti (1994) mentioned that the interpretation can vary depending on the field or context. We defined the following threshold values for interpreting the intraclass correlation values. These threshold values were constructed based on the four characteristics corresponding to the four answer options of the self-assessment scale from the first part. The numerical thresholds, on the other hand, were derived by merging the two scales, resulting in a range from 0 to 1. These are 0 (not accurate), $0.\overline{33}$ (rather not accurate), $0.\overline{66}$ (rather accurate), and 1 (completely accurate). This allows us to interpret the value of the intraclass correlation in relation to the characteristics. Calculations were performed using R (R Core Team, 2022, version 4.2.2) with the psych and irr package. To address the first question, we created an overall scale (skill 1) from each item. To address the second question, we formed the items for the various skills into scales (skills two to seven).

RESULTS

Regarding the first question, an intraclass correlation coefficient (ICC) was calculated to determine the agreement between the self-assessment and the actual performance. The analysis yielded an ICC of 0.82 ($p < .001$, 95% CI [0.76–0.86]). In relation to our concern with checking accuracy, we compare the value to the thresholds $0.\overline{66}$ and 1.

Accordingly, students assess themselves between rather accurate and completely accurate.

Regarding question two, the ICC was calculated for the different skills (see table 1). As can be seen from table 1, ICC values for skills two to six are between 0.66–0.77. For skill seven, the ICC value is lower at 0.43. Regarding the established threshold values, students can assess themselves for the skills two to six rather accurate. For skill seven, students have the tendency towards to rather not assess themselves accurately. All values are highly significant (***) indicates $p < .001$, see table 1). Overall, it was shown that the ICC value for the overall scale for fractions is higher than the average of the ICC for the different skills. We attribute this to the fact that random errors in the answers to several items are balanced out across the number, see the limits of the confidence intervals 95% CI [L, U].

skill no.	skill type	ICC	95% CI [L, U]	
skill 1	overall skill	0.815***	0.758	0.859
skill 2	add 2 fractions	0.708***	0.618	0.777
skill 3	subtract 2 fractions	0.771***	0.701	0.825
skill 4	multiply 2 fractions	0.661***	0.556	0.741
skill 5	divide 2 fractions	0.68***	0.581	0.756
skill 6	expand and contract a fraction	0.716***	0.628	0.783
skill 7	create a word problem for skills two to six	0.426***	0.248	0.562

Table 1: Overview of the results of the intraclass correlation analysis

CONCLUSIONS AND OUTLOOK

Our study has contributed to how accurate students can assess themselves in the field of fractions. We defined threshold values for the intraclass correlation coefficient values to interpret the accuracy of self-assessment. It was shown (question one) that students are between rather and completely accurate in their self-assessment overall (skill one), with a highly significant ICC value of 0.82. Upon closer inspection (question two) the accuracy of self-assessment for the skills two to six (the mastering of procedures) is lower and students' self-assessment can be interpreted as rather accurate (high significantly ICC values between 0.66 and 0.77). For the skill seven, creating word problems (the mastery of application), students tend to be rather not accurate in their self-assessments (high significantly ICC value of 0.43). So, there are differences in the accuracy of self-assessment of different types of skills in the field of fractions. We suspect that this can be conflated with the higher complexity of word problems (Pongsakdi et al., 2019; Bradshaw, 2001; Hewitt, 2005). While mastering

procedures may involve focusing on a single aspect, mastering application often requires considering multiple criteria, such as reading comprehension, interpretation, abstraction, and others.

It should be noted that our overall scale consists of five skills that require the mastery of procedures and is shaped by only one scale that involves the mastery of application. Further research should increase the number of scales for application-related tasks. Additionally, investigations could focus on how students can be supported in accurately assessing themselves, particularly for application-related tasks. In future research, we want to evaluate the data of our study more specifically. This includes examining how the accuracy is related considering further variables like the school grade and the year. We have referred to skills for the concept of fractions, which do not represent an exhaustive list of the related skills. There are other skills related to fractions, such as working with decimal numbers or interpreting fractions in various contexts, for which we cannot make any statements. Our results remain topic specific and thus cannot be transferred directly to other learning concepts. The results of our study are also primarily influenced by two schools. Accordingly, it is important to further investigate the extent to which the results can be reproduced under similar conditions or other mathematical concepts.

As mentioned at the beginning, self-regulated learning is a crucial aspect in learning mathematics. Our findings indicate that in the field of fractions, students can effectively engage in self-regulated learning phases, when focusing on procedural skills. However, applying these skills to word problems in this context, may pose a challenge for many students.

References

- Andrade, H., & Valtcheva, A. (2009). Promoting learning and achievement through self-assessment. *Theory Into Practice*, 48(1), 12–19.
<https://doi.org/10.1080/00405840802577544>
- Andrade, H. L. (2010). Students as the definitive source of formative assessment: Academic self-assessment and the self-regulation of learning. In H. L. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 106–124). Routledge.
<https://doi.org/10.4324/9780203874851>
- Behr, M. J., Lesh, R., Post, T. R., & Silver, E. A. (1983). Rational number concepts. In R. Lesh & M. Landau (Eds.), *Acquisition of mathematics concepts and processes* (pp. 91–125). Academic Press.
- Bradshaw, B. K. (2001). Do students effectively monitor their comprehension? *Reading Horizons*, 41(3), 143–154.
- Brown, G. T. L., & Harris, L. R. (2013). Student self-assessment. In J. H. McMillan (Ed.), *Sage handbook of research on classroom assessment* (pp. 367–393). Sage.
<https://doi.org/10.4135/9781452218649.n21>

- Brown, G. T. L., & Harris, L. R. (2014). The future of self-assessment in classroom practice: Reframing self-assessment as a core competency. *Frontline Learning Research*, 2(1), 22–30. <https://doi.org/10.14786/flr.v2i1.24>
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284–290. <https://doi.org/10.1037/1040-3590.6.4.284>
- De Corte, E., Verschaffel, L., & Op't Eynde, P. (2000). Self-regulation: A characteristic and a goal of mathematics education. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 687–722). Academic Press. <https://doi.org/10.1016/B978-012109890-2/50050-0>
- Hewitt, M. P. (2005). Self-evaluation accuracy among high school and middle school instrumentalists. *Journal of Research in Music Education*, 53(2), 148–161. <https://doi.org/10.1177/002242940505300205>
- Hoyle, R. H., & Dent, A. L. (2002). Developmental trajectories of skills and abilities relevant for self-regulation of learning and performance. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (pp. 49–63). Routledge. <https://doi.org/10.4324/9781315697048-4>
- Lamon, S. J. (2007). Rational numbers and proportional reasoning: Towards a theoretical framework for research. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 629–667). Information Age.
- Lortie-Forgues, H., Tian, J., & Siegler, R. S. (2015). Why is learning of fractions and decimal arithmetic so difficult? *Developmental Review*, 38, 201–221. <https://doi.org/10.1016/j.dr.2015.07.008>
- Low, R., & Jin, P. (2012). Self-regulated learning. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning* (pp. 3015–3018). Springer. https://doi.org/10.1007/978-1-4419-1428-6_202
- National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics* (Vol. 1). NCTM.
- Obersteiner, A., Reiss, K., Van Dooren, W., & Van Hoof, J. (2019). Understanding rational numbers—Obstacles for learners with and without mathematical learning difficulties. In A. Fritz-Stratmann, V. G. Haase, & P. Räsänen (Eds.), *International handbook on mathematical learning difficulties* (pp. 581–594). Springer. <https://doi.org/10.1007/978-3-319-97148-3>.
- Obersteiner, A., Dresler, T., Bieck, S. M., & Moeller, K. (2019). Understanding fractions: Integrating results from mathematics education, cognitive psychology, and neuroscience. In A. Norton & M. Alibali (Eds.), *Constructing number: Merging perspectives from psychology and mathematics education* (pp. 135–162). Springer. https://doi.org/10.1007/978-3-030-00491-0_7
- Ohlsson, S. (1988). Mathematical meaning and applicational meaning in the semantics of fractions and related concepts. In J. Hiebert & M. Behr (Eds.), *Number concepts and operations in the middle grades* (Vol. 2, pp. 53–92). Lawrence Erlbaum Associates.

- Pongsakdi, N., Kajamies, A., Veermans, K., et al. (2020). What makes mathematical word problem solving challenging? Exploring the roles of word problem characteristics, text comprehension, and arithmetic skills. *ZDM Mathematics Education*, 52, 33–44. <https://doi.org/10.1007/s11858-019-01118-9>
- Prediger, S., & Wessel, L. (2013). Fostering German-language learners' constructions of meanings for fractions—Design and effects of a language- and mathematics-integrated intervention. *Mathematics Education Research Journal*, 25(3), 435–456. <https://doi.org/10.1007/s13394-013-0079-2>
- Raaijmakers, S. F., Baars, M., Paas, F., van Merriënboer, J. J. G., & van Gog, T. (2019). Effects of self-assessment feedback on self-assessment and task-selection accuracy. *Metacognition and Learning*, 14(1), 21–42. <https://doi.org/10.1007/s11409-019-09189-5>
- Ramdass, D., & Zimmerman, B. J. (2008). Effects of self-correction strategy training on middle school students' self-efficacy, self-evaluation, and mathematics division learning. *Journal of Advanced Academics*, 20(1), 18–41. <https://doi.org/10.4219/jaa-2008-869>
- Ross, J. A., Rolheiser, C., & Hogaboam-Gray, A. (2002). Influences on student cognitions about evaluation. *Assessment in Education: Principles, Policy & Practice*, 9(1), 81–95. <https://doi.org/10.1080/09695940220119201>
- Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense-making in mathematics. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 334–370). Macmillan.
- Schunk, D. H. (1996). Goal and self-evaluative influences during children's cognitive skill learning. *American Educational Research Journal*, 33(2), 359–382. <https://doi.org/10.2307/1163289>
- Stillman, G., & Mevarech, Z. R. (2010). Metacognitive research in mathematics education: From hot topic to mature field. *ZDM: The International Journal on Mathematics Education*, 42, 145–148. <https://doi.org/10.1007/s11858-010-0245-x>
- Streiner, D. L. (2003). Starting at the beginning: An introduction to coefficient alpha and internal consistency. *Journal of Personality Assessment*, 80(1), 99–103. https://doi.org/10.1207/S15327752JPA8001_18
- Valle, C., & Andrade, H. (2014). Student self-assessment. In R. Gunstone (Ed.), *Encyclopedia of science education* (pp. 1005–1008). Springer. https://doi.org/10.1007/978-94-007-6165-0_142-4
- Zimmerman, B. J. (1998). Academic studying and the development of personal skill: A self-regulatory perspective. *Educational Psychologist*, 33(2-3), 73–86. <https://doi.org/10.1080/00461520.1998.9653292>
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, 41(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2

NOTICING OF PRESERVICE PRIMARY SCHOOL TEACHERS IN MATHEMATICS WHILE WATCHING OWN VIDEO VIGNETTES

Susanne Wöller

TUD Dresden University of Technology, Faculty of Education, Dresden, Germany

Based on the learning-to-notice framework according to Van Es and Sherin (2021), the study examines the key aspects of noticing of preservice primary school teachers in mathematics. For this purpose, a university course was developed as a study environment. Preservice teachers videorecorded their self-designed lessons with primary school children. Together with peers they reflected on their teaching observed in the video vignette through a group interview. Fifteen interviews were conducted. Data was analyzed using Grounded Theory methods (Corbin & Strauss, 2015). The results show that preservice teachers gained insights through the reflection process about various facets of children's learning processes and their own teaching activities. During the group interview they connected those facets and interpreted them in depth.

FRAMEWORK OF NOTICING

Noticing is an integral part of teachers' professional competence and can be seen as a fundamental component of teachers' expertise (Bastian et al., 2022). In this sense, Dindyal et al. (2021) aptly describe the role of noticing in the context of teachers' everyday demands in classroom situations:

Every act of teaching depends on noticing: noticing what children are doing, how they respond, evaluating what is being said or done against expectations and criteria, and considering what might be said or done next. (p. 1)

In a nutshell, this is about perceiving, evaluating and reflecting on current and for future teaching activities. According to Stahnke et al. (2016), the theoretical construct of noticing is defined differently depending on the literature. Some authors consider this to mean only the perception of relevant situations. Others summarize the perception and interpretation of these situations. Still others define noticing more broadly and subsume perception, interpretation, and decision-making for further practice.

For this study, I refer to the revised learning-to-notice framework according to Van Es and Sherin (2021). In one of their first publications, Van Es and Sherin (2002) formulated key aspects of noticing, from which the learning-to-notice framework had developed. They proposed a methodological tool of using video vignettes for teacher education. Inservice teachers watched video clips of their own lessons as part of a course at the university. They then analyzed these clips regarding children's thinking, teachers' roles and discourse. From the results of the study Van Es and Sherin (2002) derived the key aspects of noticing: First, teachers identify important or noteworthy classroom situations. That means they observe what children are doing or saying and

how they are thinking about a subject matter. Then, teachers make connections between the specifics of classroom interaction and the broader principles of teaching and learning. By doing so, they use their knowledge of the context to reason about classroom interactions. These processes are very closely linked, whereas *interpretation* plays an important role in the development of noticing (Van Es & Sherin, 2002).

Later, they included the terms *attending*, *interpreting*, and *shaping* to their framework (Van Es & Sherin, 2021): Teachers identify noteworthy features of classroom interaction and disregard selected aspects that are less consequential (*attending* with reference to *selective attention*). They use their knowledge and experiences to make sense of what they observed and adopt a stance of inquiry (*interpreting*). Additionally, teachers construct interactions and contexts that provide access to additional information of children's thinking. In the video they actively seek additional information that could become object of further attending and interpreting (*shaping*).

The related literature includes some similar pre-existing terms. Thus, Van Es and Sherin (2008) themselves draw references to the term *teacher reflection* by writing "learning to notice is one important dimension within the process of reflection" (p. 247). In this respect, teachers' reflection enables them to make sense of their experiences of classroom situations and use this knowledge for future pedagogical decisions (Van Es & Sherin, 2008). Relationships to the term *analyzing* are also visible. Kuntze et al. (2015) wrote that analyzing teaching situations involves identifying relevant information, critically evaluating this information against the background of underlying theories, and finally articulating these examinations.

Obviously, noticing is a part of reflection and analyzing, but the respective terms are defined even more broadly. The revised learning-to-notice framework by Van Es and Sherin (2021) reflects the complex process very well: Noticing is not just the mere perception of relevant situations, but the complex interplay of perceiving, interpreting, searching for information, and thus perceiving new relevant situations.

VIDEO VIGNETTES AS A TOOL IN TEACHER EDUCATION

Videos can be used as *artifacts of practice* to support noticing processes (Dindyal et al., 2021). For this reason, many studies that investigate the noticing of preservice or experienced teachers use video vignettes as the preferred tool. This recorded excerpts of varying lengths from own or other teachers' lessons are used to trigger reflection processes among the observers (Gaudin & Chaliès, 2015). Given prompts indicate which aspects of the video should be analyzed. These prompts vary from detailed questions to more open suggestions. Typically, short, edited segments of videos are used to engage in dialogue with the observer (Star & Strickland, 2008).

The use of video vignettes reveals various advantages, especially in teacher education. Authentic situations of children's learning and teachers' activities are much better represented in video recordings than in observation protocols or similar text-based materials (Bastian et al., 2022). In addition, rich and complex teaching situations are

slowed down, so that segments of the lesson can be analyzed more precisely than when they are recalled from memory. This allows them to focus on relevant situations or children's solutions, and even look at those several times with different perspectives (Van Es & Sherin, 2002). Another advantage is that video excerpts can be discussed in groups with other teachers or lecturers (Van Es & Sherin, 2008). In fact, the use of video vignettes of one's own lessons offers advantages. Thus, details of the classroom teaching that may not have been noticed during the lesson can be recognized (Star & Strickland, 2008). The most prominent examples of this are children's interactions and conversations during group phases.

Some studies examined the noticing of preservice teachers using videos. In general, these studies show that the key aspects of noticing in preservice teachers are less developed than in experienced teachers (Gaudin & Chaliès, 2015). Preservice teachers mostly have difficulties interpreting children's working processes, solutions and errors (Stahnke et al., 2016). Moreover, they are often unable to offer well-founded reasons for their reflections. This means that they tend to remain at a superficial level on what is happening in the classroom (Ulusoy, 2020). Furthermore, they are more likely to "view a lesson merely as a chronological but disconnected sequence of events" (Star & Strickland, 2008, p. 110) than experienced teachers do.

FOCUS OF THE STUDY AND RESEARCH QUESTIONS

On the one hand, the aim of this qualitative study is to identify and differentiate the possible dimensions of the key aspects of noticing in preservice teachers mentioned by Van Es and Sherin (2021). On the other hand, teacher education is to be improved by offering preservice teachers the opportunity to reflect on their own teaching experiences as part of a university course. The aim is to interlink theory and practice more closely during teacher education programs. In this way, the knowledge gained by preservice teachers should be as high and sustainable as possible to prepare them for their professional careers as a teacher and to enhance their professional development.

Therefore, in this study, several promising approaches were combined to foster preservice mathematics teachers noticing: As part of a specially designed university course preservice teachers worked in groups with their peers. They planned mathematics teaching and reflected on their self-conducted lessons collectively. To do so, they used their own video vignette as an artefact during a group interview. In this context, it is interesting to examine what exactly the preservice teachers retrospectively pay attention to in their own mathematics teaching situations, shortly before starting their professional careers. Based on the key aspects of Van Es and Sherin's framework (2021), the study addresses the following overarching and subordinate research questions: *What key aspects of noticing become apparent in a group interview of preservice mathematics teachers while watching their own video vignette?*

RQ1: *What do preservice teachers retrospectively attend in classroom situations?*

RQ2: *How do preservice teachers retrospectively interpret the teaching situations?*

METHODS

The participants of the study took part in a 15-week course at a German University, which I taught myself. As part of the course the preservice teachers worked in groups of two to three on a design of open tasks for math lessons in primary school (grade 1 to 4). To deal with the heterogeneity of primary school children, open tasks play an important role in teacher education programs in Germany. In the course the participants learned about the characteristics of open tasks and how they can be implemented. After that the groups taught the designed lessons themselves in primary school classes. Those lessons were videorecorded. This resulted in video vignettes with an average length of 25 minutes per group. These vignettes showed both the children's working on the task and the teaching activities of the preservice teachers. Both a detailed description of my course, and a detailed description how and under which conditions the video vignettes were created, are described in Wöller (2023).

I ran this course twice, once in winter 2022/23 and then in 2023/24. In total, 32 preservice primary school teachers took part in the study. 30 were female and two were male. The average age was 23. The participants were in the 7th semester of their teacher training program for primary schools. This means they have already learnt basic information about the didactics of mathematics, subject-specific knowledge in mathematics and pedagogy. Apart from the usual four-week internships at primary schools, they had no other teaching experience. They were about to finish their studies and start their mandatory internship (*Referendariat*) at primary school in Germany.

After the preservice teachers held their lessons, I conducted interviews with each group. A key element of the interview was watching the video vignette together. I gave an open prompt: *Pause the video if you notice something you'd like to talk about. I am interested in your thoughts about children's learning processes and your own teaching activities.* During the interview, a screen recording of the video vignette was made, and the participants' voices were recorded. A total of 15 interviews were conducted over two survey periods. On average, the interviews lasted 1 hour per group. Transcripts of the group interviews were created using *f4* software. At the points where the preservice teachers paused the video to talk about, screenshots were made and inserted into the transcript. All transcripts, including the screenshots, were transferred to *ATLAS.ti* and analyzed using Grounded Theory methods (Corbin & Strauss, 2015).

In the process of open coding, I first discovered phenomena in the data material without theory biased assumptions in mind. Among other things, I explored whether and to what extent 'recurring patterns' in the participants' reflection processes can be discovered: Which situations caused the preservice teachers to pause the video? How deeply did they analyze these situations? Did they resume and continue their analysis with subsequent interesting video sequences? ...

In the process of axial coding, the discovered phenomena were placed in relation to each other. Memos and concept maps helped to build categories and uncover

relationships between them. In this way, the categories were increasingly abstracted. The next step will be the process of selective coding, from which point a theory grounded in empirical data will emerge. This step is expected to lead to a further differentiation of the learning-to-notice framework based on Van Es and Sherin (2021).

PRELIMINARY RESULTS

After the initial analysis, it is obvious that the results are incredibly exciting and promising. I can only provide insights into selected phenomena and categories here. To visualize some results, I included transcript excerpts from my data material translated into English. These excerpts are exemplary for many other interview sequences in which similar phenomena were revealed. Along with the two subordinate research questions, I will first illustrate *what* preservice teachers *attend* in teaching situations when watching their own video vignette in the group. In the second step, I will show *how* they *interpret* these situations in retrospect. It is important to mention that RQ1 and RQ2 are closely linked and cannot be answered without overlapping.

In their reflections, the participants focused on various actors in the teaching situation seen in the video. On the one hand, they observed the group of learners. In doing so, they focused on the children's individual learning processes and included learning progress that occurred during the lesson. Among other things, they looked at how the children developed mathematical concepts, or which difficulties emerged during the lesson. Moreover, the participants considered the complex structure of interaction in group phases and the dynamics of the classroom situation. The following transcript excerpt (from a lesson on creating and sorting snowflake shapes) illustrates this:

Anna But here you can see that only the scribe really has something to do in the moment. The others are all a bit, yes, looking a bit at what's happening, even though one girl has the worksheet in her hand and there are still a few tasks to do, such as sorting the snowflakes. But they haven't really got to work on that yet. They're waiting until one thing is finished and then they all go to the next step together. (2_4_10012024, #00:13:22)

On the other hand, the participants observed their own teaching actions and previous methodical decisions. Among others, they looked at how they instructed teaching phases and where they intervened in children's learning processes. Moreover, they reflected on their own behavior in front of the class or their language, i.e. how they phrased tasks or provided explanations when the children asked questions. In doing so, they considered whether their actions were helpful for the children's learning or maybe a different approach might be more promising in the future. The following transcript excerpt (from a lesson on producing all possible square quadruplets) illustrates this:

Martin Here, in retrospect, when I see the explanation, I don't know what, how or why I did it that way, that I first talk about the reflection and say that with the symmetry axis. They even figured it out on their own. And then, I continue with the rotation. Yes, of course, it's connected somehow, but looking back, it was totally incoherent. (2_3_10012024, #00:34:27)

In this sequence, the participant noticed in retrospect that his explanations in front of the class were not effective, as the children had already worked out these mathematical relationships in the previous group phase. Consequently, his discourse did not address the current needs of the children, but rather seemed to confuse them. Obviously, he critically evaluated his own teaching actions in relation to the children's current level of mathematical thinking and understanding.

This phenomenon is not an individual case in the data material but can be observed in most of the participants in the study. In doing so, preservice teachers analyzed the effects of their teaching and interpreted children's learning processes. For their explanations during the interview, among others, they used their pedagogical, didactic and subject-specific knowledge to argue about how good math lessons should be designed and how the children in the class can be supported in their (mathematical) learning processes. The following transcript excerpt (again on snowflake shapes), in which the participant finally reflected on her lesson, visualizes this phenomenon:

Lena I would do the same thing again so that the children can try things out for themselves. If they do something by themselves, they better understand these mathematical topics. [...] What I would do differently is to give them other help. We did have some children who didn't figure it out or who only got there through the hint cards. That you somehow guide them a bit more without taking away their freedom. (1_6_11012023, #01:04:36)

In addition, the participants actively searched for further information in the video to support their interpretations, especially in children's written documentations or their statements. For that they placed the video sequences in relation to each other. The following transcript excerpt (from a lesson on folding cuts) illustrates how deeply the participants tried to empathize with the children's thoughts and looked for sequences in the video that they can use for their interpretations:

Sophie What I notice with her is that she also drew the edge at the very beginning, and since then she has orientated herself on one side of her page and I somehow have the feeling that she has understood, 'Okay, good, that's my axis now', I would say. 'And if I just draw on one side, then it works across'. That she's already understood the principle that you always draw on one side, not like everyone else, I'll say. And then she can create what she really wants by folding it out. So that she can't be surprised at all because she knows what she wants to do. (1_8_18012023, #00:18:45)

The preservice teachers often concluded their interpretation of a perceived sequence with well-founded thoughts on alternative or future teaching activities. They then either attended new meaningful situations that they interpret, or they collected further arguments for their previous interpretation processes through subsequent sequences.

DISCUSSION AND PERSPECTIVES

The results clearly showed that the preservice teachers in the study did not reflect on their teaching actions superficially at a descriptive level. They considered the complex

contexts of the teaching situations depicted in the video vignette. They interpreted children's learning processes in depth as well as related this process to their own teaching activities. In doing so, they used their existing knowledge and actively searched for evidence to support or differentiate their considerations about children's mathematical understanding. Consequently, a complex network of interpretations rather than isolated sequences of descriptions was formed during the conversation of preservice teachers while watching their own video vignette. The results of the study thus enlighten the key aspects of teacher noticing defined by Van Es and Sherin (2021) and show how *attending*, *interpreting* and *shaping* are evident in preservice teachers who took part in the study.

Surprisingly, the results of this study contradict the results of previous studies (cf. Stahnke et al., 2016; Star & Strickland, 2008). There are various possible reasons for this. It is conceivable that long-term engagement with a mathematical topic as part of the university course could have an influence on the noticing of preservice teachers in this study environment. After all, the course lasted a total of 15 weeks. This is a lot of time to devote to mathematical content and didactic preparation. Moreover, preservice teachers reflected on their teaching *together* with 'like-minded people' during the group interview – in other words, with people who have similar expectations, backgrounds and problems to oneself. Thus, Van Es and Sherin (2008) admit that this study environment can improve the key aspects of noticing of teachers, especially when own video vignettes were used.

Nonetheless, some aspects of the study must be evaluated critically. One of the main points is that the preservice teachers attended the university course with me and were then also interviewed by myself. In addition, I am also the coder of the data material. This study environment inevitably leads to a subjectively biased discovery of phenomena during the coding process. Another important point is that the preservice teachers took an examination as part of the course. This could put them under pressure to perform well during the interview. Nevertheless, the results emphasize the relevance of collective reflection phases in teacher education through the sensible use of methods such as self-created video vignettes. The video as an artifact gives preservice teachers the opportunity to learn from their own teaching experiences and to intensively analyze the learning processes of the children. This approach should therefore be incorporated into teacher education programs to foster preservice teachers noticing, and thus further advance the interlinking of theory and practice.

In the next step, the data is further axially and selectively coded to differentiate the learning-to-notice framework according to Van Es and Sherin (2021). As mentioned above, I am still in the process of analyzing the data. Besides, additional group interviews with preservice teachers were conducted in winter 2024/25. For this, the course was adapted and adjusted to the content-related needs of preservice teachers. The effects of this will be considered in future research.

References

- Bastian, A., Kaiser, G., Meyer, D., Schwarz, B. & König, J. (2022). Teacher noticing and its growth toward expertise: An expert–novice comparison with pre-service and in-service secondary mathematics teachers. *Educational Studies in Mathematics*, 110(2), 205-232. <https://doi.org/10.1007/s10649-021-10128-y>
- Corbin, J. M. & Strauss, A. L. (2015). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Sage publications, inc.
- Dindyal, J., Schack, E. O., Choy, B. H. & Sherin, M. G. (2021). Exploring the terrains of mathematics teacher noticing. *ZDM–Mathematics Education*, 53(1), 1-16. <https://doi.org/10.1007/s11858-021-01249-y>
- Gaudin, C. & Chaliès, S. (2015). Video viewing in teacher education and professional development: A literature review. *Educational Research Review*, 16, 41-67. <https://doi.org/10.1016/j.edurev.2015.06.001>
- Kuntze, S., Dreher, A. & Friesen, M. (2015). Teachers’ resources in analyzing mathematical content and classroom situations: The case of using multiple representations. In K. Krainer & N. Vondrová (Eds.), *Proceedings of the Ninth Congress of the European Society for Research in Mathematics Education (CERME9)* (pp. 3213-3219). Charles University in Prague, Faculty of Education and ERME. <https://hal.science/hal-01289863>
- Stahnke, R., Schueler, S. & Roesken-Winter, B. (2016). Teachers’ perception, interpretation, and decision-making: A systematic review of empirical mathematics education research. *ZDM–Mathematics Education*, 48, 1-27. <https://doi.org/10.1007/s11858-016-0775-y>
- Star, J. R. & Strickland, S. K. (2008). Learning to observe: Using video to improve preservice mathematics teachers’ ability to notice. *Journal of Mathematics Teacher Education*, 11, 107-125. <https://doi.org/10.1007/s10857-007-9063-7>
- Ulusoy, F. (2020). Prospective teachers’ skills of attending, interpreting and responding to content-specific characteristics of mathematics instruction in classroom videos. *Teaching and Teacher Education*, 94, 1-13. <https://doi.org/10.1016/j.tate.2020.103103>
- Van Es, E. A. & Sherin, M. G. (2002). Learning to notice: Scaffolding new teachers’ interpretations of classroom interactions. *Journal of Technology and Teacher Education*, 10(4), 571-596.
- Van Es, E. A. & Sherin, M. G. (2008). Mathematics teachers’ “learning to notice” in the context of a video club. *Teaching and Teacher Education*, 24(2), 244-276. <https://doi.org/10.1016/j.tate.2006.11.005>
- Van Es, E. A. & Sherin, M. G. (2021). Expanding on prior conceptualizations of teacher noticing. *ZDM–Mathematics Education*, 53, 17-27. <https://doi.org/10.1007/s11858-020-01211-4>
- Wöller, S. (2023). Professional Reflective Competence of Prospective Primary School Teachers in Mathematics Based on Video Vignettes. In P. Drijvers, H. Palmér, C. Csapodi, K. Gosztonyi & E. Kónya (Eds.), *Proceedings of the Thirteenth Congress of the European Society for Research in Mathematics Education (CERME13)* (pp. 715-722). Alfréd Rényi Institute of Mathematics and ERME. <https://hal.science/hal-04421650>

CAN GERMANY LEARN FROM “DIGITAL DENMARK”? DO THE DIGITALIZATION-RELATED ATTITUDES OF GERMAN AND DANISH TEACHERS DIFFER?

Mira H. Wulff, Anika Radkowitsch & Aiso Heinze

IPN Leibniz Institute for Science and Mathematics Education, Kiel (Germany)

Germany's level of digitalization in schools is below standard, both in comparison to the global and European averages. Conversely, Denmark, a neighboring nation, is frequently cited as a paradigm of digitalization in schools. This study aims to explore the question of whether there is a discrepancy in attitudes concerning the use of digital technology between mathematics teachers in Germany and Denmark. A reanalysis of data from two surveys of teachers with $N = 155$ and $N = 68$ teachers from the German-Danish border region revealed that German teachers generally ascribe a higher task value to the approach, while Danish teachers had higher self-efficacy regarding the approach. The results of the study indicate that the provision of resources is a pertinent factor in the implementation of up-to-date mathematics teaching.

INTRODUCTION

Increasing digitalization has led to a growing focus on the incorporation of digital technologies in mathematics teaching in primary and secondary schools. Teachers play a pivotal role in this implementation because they are responsible for the design, execution, and assessment of the lessons. Consequently, facets of digitization-related knowledge, as well as teachers' attitudes and beliefs concerning teaching and learning with digital technologies, have gained significance. These include teachers' *self-efficacy* in technology-related knowledge domains and the three value categories of *task value beliefs* (*intrinsic, attainment and utility value*; see Wigfield et al., 1992, 1994), as well as pedagogical, subject-specific, and didactic knowledge (e.g., Fernández-Batanero et al., 2020). The initial and secondary phases of teacher training (five years of study and one and a half years of traineeship) serve as the foundational preparation for educators. However, the third phase of teacher training, known as in-service training, is gaining prominence due to the perpetual evolution of digital technologies and their applications in the classroom. Further training opportunities allow educators to acquaint themselves with novel teaching media and to further develop their understanding of digitalization. The participation in this phase of training is voluntary and determined by the individual teacher's interests and needs, contingent on the specific policies of the respective federal state. Given the competition with other in-service teacher training offers, research into the factors determining participation is of particular interest.

A socio-cognitive learning theory approach (Bandura, 1986) that transcends national boundaries can provide a valuable contribution to the development of appealing further

education programs. When assessing the degree of digitalization in German schools, a useful comparison can be made with its neighboring country Denmark. According to the ICILS 2023 findings, Denmark's level of digitalization in schools is considerably more advanced than Germany's (Fraillon, 2024). In this context, a particularly salient comparison can be drawn between the neighboring provinces Schleswig-Holstein (Northern Germany) and Syddanmark (Southern Denmark). These two provinces show only minor cultural differences, attributable to the shared history of the area, which has been influenced by several significant border shifts over the last two centuries. These historical shifts in the German-Danish border have led to the presence of notable ethnic minorities on both sides of the border, which have maintained their political and cultural identities, contributing to a complex and dynamic social landscape. These manifest themselves, for example, in their own minority schools or in exemptions for political participation in regional and national parliaments in Germany and Denmark. These minor cultural differences make the two provinces a fruitful field of study for investigating factors for the considerable differences in the degree of digitalization of schools in Germany and Denmark.

THEORETICAL BACKGROUND

The importance of integrating digital media into pedagogical practices is a subject that has gained increasing traction in recent years. This trend is evidenced by the adoption of digital media in educational institutions, as evidenced by the work of the OECD (2020). A primary rationale for this emphasis is the objective of equipping students with the competencies necessary to thrive in the digitalized workplace. Ensuring the effective integration of digital media in teaching necessitates the consideration of numerous factors. On the one hand, institutional framework conditions must be considered, such as the equipment of schools with digital technologies. Research indicates a positive correlation between the availability of these technologies and the perceived relevance of integrating digital technologies by teachers, as well as their willingness to engage with them (Petko et al., 2018). On the other hand, the respective subject teacher is responsible for making methodological and didactic decisions in the planning and implementation of subject lessons.

A variety of conditions for success have been identified as integral to the integration of digital technologies in the classroom. According to Ertmer (1999), these conditions can be categorized into external and internal barriers. Externally, barriers such as hardware and software, teacher training, and didactic and technical support are implicated. Due to the gradual integration of digital technologies in the initial training of teachers, as well as the increasing digitalization of schools, it is plausible to assume that external barriers will be superseded over time. Consequently, these barriers can be regarded as temporary challenges for the integration of digital technologies.

Internal barriers in the use of digital technologies in the classroom

The internal barriers encompass a variety of factors, including, but not limited to, the teacher's unconscious attitudes regarding teaching and learning with digital

technology. These attitudes may encompass beliefs about digital technology, as well as the teacher's professional knowledge (e.g., TPACK; Mishra & Koehler, 2006) and corresponding *self-efficacy*. These internal barriers are often considered the "biggest enablers" (Ertmer et al., 2012) when compared to external barriers, as they are more comprehensive and relevant in the process of integrating digital technologies into subject teaching (Ertmer, 1999). Consequently, the present study prioritizes an examination of internal barriers.

The integration of innovative pedagogical approaches into mathematics teaching is contingent upon the alignment between these approaches and the teacher's prevailing beliefs (Zhao & Frank, 2003). In accordance with the expectation-value theory proposed by Wigfield et al. (1994), the *task value beliefs* of teachers can be regarded as the alignment between their personal beliefs and the pedagogical approach with digital technology within mathematics lessons. According to Wigfield et al. (1992, 1994), the *task value beliefs* cover three value categories: *intrinsic value*, *attainment value*, and *utility value*: (1) The *intrinsic value* category is defined as the teacher's subjective enjoyment of the pedagogical approach, (2) the category *attainment value* comprises the teacher's individually perceived relevance of the approach and the associated ability to master the pedagogical approach and the associated challenges in preparation and implementation, and finally, (3) the *utility value* category is defined as the perceived individual benefit of the pedagogical approach for the teacher's own goals as well as their overarching (teaching) goals. The *task value* attributed to a pedagogical approach, in conjunction with the expectation of success assigned to said approach, has been demonstrated to be a significant predictor of an individual's achievement motivation (see Wigfield & Eccles, 1992). Therefore, it can be hypothesized that teachers will adopt a new pedagogical approach with digital technology in their regular lessons if they assign a high *task value* to it.

In addition to the ascribed *task value*, the individual knowledge of the teacher also has a significant influence on the decision as to how a lesson is carried out (Borko & Putnam, 1995). The concept of teacher knowledge can be delineated through various theoretical frameworks. For instance, the TPACK framework, as developed by Mishra and Koehler (2006) and others, builds upon the seminal PCK model by Shulman (1986) by incorporating technological knowledge (TK) and its associated overlaps, namely technological pedagogical knowledge (TPK), technological content knowledge (TCK), and technological pedagogical content knowledge (TPCK). This addition to the PCK model is predicated on the assumption that the interaction of technology, content and pedagogy is fundamental to the effective integration of technology in subject teaching (Angeli & Valanidis, 2009). In addition to teachers' knowledge in relation to the aforementioned areas, their *self-efficacy* in relation to these areas are also relevant (e.g., Wozney et al., 2006). The concept of *self-efficacy* in relation to a knowledge construct signifies the individual assessment of an educator's capacity and efficacy in acting in relation to this knowledge construct. A body of extant research indicates a positive correlation between the use (or intention to use) of digital technologies in the classroom

and *self-efficacy* with regard to technology-related knowledge constructs (Wozney et al., 2006; Wulff et al., 2023a). According to Bauer and Kenton (2005), the level of *self-efficacy* of teachers is even more relevant for the (intention to) use digital technologies in the classroom than the level of knowledge or skills of the teachers.

RESEARCH QUESTION

Based on the presented framework, we examine the following research question. To what extent do German and Danish teachers from the German-Danish border provinces differ in terms of the extent of their internal barriers (i.e., *task value beliefs* and *self-efficacy* regarding TK, TPK and TPACK) to the use of digital technologies?

METHOD

In order to investigate this question, data collected as part of a binational educational design research project was subjected to a secondary analysis. The project's primary focus was on the utilization of 3D printing technology as a learning context within mathematics lessons (cf. Wulff et al., 2023b). The topic of 3D printing is particularly suitable for the study, as it is not a standard digital topic for teachers in mathematics lessons. Data was collected using an online questionnaire from teachers in the German-Danish border provinces Schleswig-Holstein (Northern Germany) and Syddanmark (Southern Denmark) in two different surveys (excluding teachers from German and Danish minority schools). The first survey concentrated on the prevailing conditions and framework for the integration of 3D printing technology into teaching. To this end, 155 teachers were surveyed, including 113 teachers from schools in Germany and 42 teachers from schools in Denmark (cf. Wulff et al., 2023a). The second survey was conducted after registration for but before a short training course for mathematics teachers on the use of 3D printing technology as a learning context in regular mathematics lessons. The 68 participants included 39 teachers from schools in Germany and 28 teachers from schools in Denmark. The questionnaire included scales on the teachers' *self-efficacy* regarding the technology-linked knowledge areas *TK*, *TPK*, and *TPACK* as well as scales on their *task value beliefs* regarding 3D printing technology distinguishing intrinsic value, attainment value, and utility value (all scales with 4-point Likert scale 1 = disagree to 4 = agree; for details see Wulff et al., 2023a and Tab. 1).

Scale (# items)	Sample item	Reliability
Intrinsic value (3)	From a teacher's perspective, I find integrating 3D printing technology into my math lessons interesting.	.82/.77
Attainment value (2)	It is important to me to be good at integrating 3D printing technology into my math lessons.	.81/.71
Utility value (2)	3D printing technology is useful/helpful in math classrooms to enhance learning opportunities for students.	.77/.73

Self-efficacy TK (4)	I am convinced that I can solve technical problems in the use of 3D printing technology.	.88/.86
Self-efficacy TPK (4)	I am convinced that I can use 3D printing technology to enhance student learning in a math lesson.	.73/.82
Self-efficacy TPACK (4)	I am convinced that I can integrate 3D printing technology appropriately into mathematics lessons and the associated methods.	.80/.78

Table 1: Information on questionnaire scales; reliability for Survey 1 and 2 based on Cronbach's α (scales with ≥ 3 items) and Spearman Brown (scales with 2 items).

RESULTS

	Survey 1		Survey 2	
	Teachers from Germany	Teachers from Denmark	Teachers from Germany	Teachers from Denmark
Intrinsic value	3,39 (0,60)*	3,10 (0,79)	3,48 (0,44)	3,53 (0,48)
Attainment value	2,96 (0,90)**	2,49 (1,11)	3,23 (0,88)	3,05 (0,88)
Utility value	3,18 (0,64)*	2,89 (0,90)	3,24 (0,51)	3,52 (0,45)*
Self-efficacy TK	3,02 (0,72)	3,09 (0,64)	2,96 (0,67)	3,09 (0,53)
Self-efficacy TPK	3,10 (0,62)[#]	2,94 (0,66)	3,15 (0,50)	3,19 (0,52)
Self-efficacy TPACK	2,91 (0,65)	2,75 (0,72)	3,00 (0,67)	3,08 (0,49)

Table 2: Binational comparison: mean values (standard deviations) and t -test-results for the two surveys with German and Danish teachers (** $p < .01$, * $p < .05$, [#] $p < .10$).

The data analysis of the collected data indicates that German teachers ascribe a considerably higher *task value* to the implementation of 3D printing technology in mathematics lessons across all three value categories (*intrinsic value*, *attainment value*, *utility value*) than Danish teachers. Additionally, they exhibit a heightened level of *self-efficacy* related to *TPK* (see Tab. 2, Survey 1). In the Survey 2 sample of teachers who intend to participate in 3D printing-related training, Danish teachers attribute a significantly higher *utility value* to the use of the technology (see Tab. 2, Survey 2).

A national comparison of the data from the two surveys for the teachers from Germany demonstrates that teachers who plan to participate in the training courses (Survey 2) tend to ascribe a greater *attainment value* to the use of 3D printing technology in mathematics lessons (see Tab. 3, Teachers from Germany). A similar national

comparison for teachers from Denmark also demonstrates that teachers expressing an interest in the training ascribe a significantly higher *task value* to the use of the technology (in all three value categories) in comparison to those from Survey 1. Furthermore, they exhibit a heightened level of *self-efficacy* related to *TPK* and *TPACK* (see Tab. 3, Teachers from Denmark).

	Teachers from Germany		Teachers from Denmark	
	Survey 1	Survey 2	Survey 1	Survey 2
Intrinsic value	3,39 (0,60)	3,48 (0,44)	3,10 (0,79)	3,53 (0,48)*
Attainment value	2,96 (0,90)	3,23 (0,88)#	2,49 (1,11)	3,05 (0,88)*
Utility value	3,18 (0,64)	3,24 (0,51)	2,89 (0,90)	3,52 (0,45)**
Self-efficacy TK	3,02 (0,72)	2,96 (0,67)	3,09 (0,64)	3,09 (0,53)
Self-efficacy TPK	3,10 (0,62)	3,15 (0,50)	2,94 (0,66)	3,19 (0,52)*
Self-efficacy TPACK	2,91 (0,65)	3,00 (0,67)	2,75 (0,72)	3,08 (0,49)#

Table 3: National comparison: mean values (standard deviations) and *t*-test-results for (a) the German and (b) the Danish subsamples within the two surveys (** $p < .01$, * $p < .05$, # $p < .10$).

DISCUSSION

The data indicates that teachers who plan to participate in a 3D printing-related training, irrespective of their nationality, attribute higher or similar *task values* to the technology than teachers from Survey 1 (Tab. 3). In the context of *self-efficacy* related to the three technology-related knowledge domains, teachers from Survey 2 demonstrated higher or similar values in comparison to teachers from Survey 1. It seems that teachers' prevailing beliefs and *self-efficacy* exert an influence on their selection of training. The data also show that German teachers from the general Survey 1 attribute higher *task values* to the use of 3D printing technology in mathematics lessons than Danish teachers (Tab. 2). This tendency may be attributed to the earlier integration of 3D printing technology into Danish school education (cf. Fraillon, 2024) and, consequently, a higher experience of Danish teachers in integrating the technology in their mathematics lessons. Therefore, it can hypothesized that Danish teachers have adapted their beliefs to reality and that teachers only register for the training if they are committed to incorporating 3D printing technology in mathematics lessons based on their personal beliefs. In Germany, the integration of 3D printing technology into school curriculum has been limited to a few years, suggesting that teachers may not have had sufficient opportunity to develop a comprehensive understanding of and beliefs regarding the technology and the skills to apply it in diverse and sophisticated ways. The question that remains is whether the internal barriers can be categorized as the “biggest enablers” (Ertmer et al., 2021) in Germany's delayed implementation of new digital technologies in mathematics education. The data presented in this study do

not support this hypothesis. Consequently, the study's findings suggest that external barriers may be the primary factor contributing to Germany's lag in adopting new digital technologies. The study indicates that teachers' beliefs regarding the implementation of digital technologies (Zhao & Frank, 2003) in Germany are comparable to those in Denmark, suggesting that the availability of digital and technological resources is a contributing factor to the observed differences in implementation.

The present study is subject to several limitations. For instance, existing data from the German-Danish border provinces Schleswig-Holstein and Syddanmark were used for the present study. Second, both samples are presumably selective from another point of view: While all schools in both regions were contacted in Survey 1, the topic of the survey (integrating 3D printing technology in mathematics lessons) was clearly articulated from the outset. Consequently, it is plausible that only teachers with a favorable disposition towards the technology participated in this survey. Third, only a restricted set of variables with only a few items were administered in the two surveys.

Despite the limitations, the results of this study suggest an interesting conclusion: It is not necessarily teachers' attitudes in terms of internal barriers that hinder the integration of new digital technologies in mathematics education. The integration of new digital technologies can be facilitated by the provision of resources, including hard- and software, as well as teacher training, as in the example of Denmark and the advanced 3D printing technology. This would enable a significant proportion of mathematics teachers to utilize this technology, thereby offering students a more contemporary mathematics education.

References

- Angeli, C., & Valanides, N. (2009). Epistemological and method-ological issues for the conceptualization, development, and assessment of ICT-TPCK: Advances in technological pedagogical content knowledge (TPCK). *Computers & Education*, 52, 154-168.
- Bandura, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive Theory*. Prentice-Hall.
- Borko, H., & Putnam, R. T. (1995). Expanding a teacher's knowledge base: A cognitive psychological perspective on professional development. In T. R. Guskey, & M. Huberman (Eds.), *Professional development in education: New paradigms and practices* (pp. 35-65). Teachers College Press.
- Ertmer, P. A. (1999). Addressing first- and second-order barriers to change: Strategies for technology integration. *Educational Technology Research and Development*, 47, 47-61.
- Ertmer, P. A., Ottenbreit-Leftwich, A. T., Sadik, O., Sendurur, E., & Sendurur, P. (2012). Teacher beliefs and technology integration practices: A critical relationship. *Computers & Education*, 59(2), 423-435.

- Fernández-Batanero, J. M., Montenegro-Rueda, M., Fernández-Cerero, J., & García-Martínez, I. (2020). Digital competences for teacher professional development. Systematic review. *European Journal of Teacher Education*, 45(4), 513–531. doi: 10.1080/02619768.2020.1827389
- Fraillon, J. (Ed.) (2024). *An International Perspective on Digital Literacy. Results from ICILS 2023*. IEA.
- Mishra, P. & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, 108 (6), 1017-1054.
- OECD (2020), *Back to the Future of Education: Four OECD Scenarios for Schooling*. Educational Research and Innovation, OECD Publishing, Paris, doi: 10.1787/178ef527-en
- Petko, D., Prasse, D., & Cantieni, A. (2018). The interplay of school readiness and teacher readiness for educational technology integration: A structural equation model. *Computers in the Schools*, 1-18.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15 (2), 4-14.
- Wigfield, A., & Eccles, J. S. (1992). The development of achievement task values: A theoretical analysis. *Developmental Review*, 12(3), 265–310.
- Wigfield, A. (1994). Expectancy-value theory of achievement motivation: A developmental perspective. *Educational Psychology Review*, 6(1), 49-78.
- Wulff, M. H., Radkowitsch, A., Wilken, M., & Heinze, A. (2023a). Wie sehen Lehrkräfte die Nutzung des 3D-Drucks als Lern-kontext im Mathematikunterricht der Sekundarstufe? In F. Dilling, D. Thurm, & I. Witzke (Eds.), *Digitaler Mathematik-unterricht in Forschung und Praxis* (pp. 263-273). WTM. doi: 10.37626/GA9783959872041.0.26
- Wulff, M. H., Wilken, M., & Heinze, A. (2023b). Increasing the skills on occupationally relevant digital technologies among students in Southern Denmark and Northern Germany: 3D printing as a learning context in regular mathematics class. In F. Dilling, F. Pielsticker, & I. Witzke (Eds.), *Learning mathematics in the context of 3D printing* (pp. 51-72). Springer Spektrum. doi: 1007/978-3-658-38867-6_3
- Wozney, L., Venkatesh, V., & Abrami, P. (2006). Implementing Computer Technologies: Teachers' Perceptions and Practices. *Journal of Information Technology for Teacher Education*, 14, 173-207.
- Zhao, Y., & Frank, K. A., (2003). An ecological analysis of factors affecting technology use in schools. *American Educational Research Journal*, 40, 807-840.

UNEXPECTED RICHNESS FROM A WOODEN ARCH PUZZLE: EMBODIED WAYS OF LEARNING ABOUT PARABOLAS

Paul Woo Dong Yu¹, Joy A. Oslund¹, and Tasnim A. Alshuli²

¹ Grand Valley State University (U.S.A), ²University of Arizona (U.S.A)

This case study confronts ableist notions of mathematics. Concepts of embodied cognition were used in the classroom to teach parabolas to a blind student using a wooden puzzle. To accommodate the student's learning, a tactile game-based activity was used to explore foundational parabola concepts. What was intended to be a basic introduction to quadratic parabolas evolved into an expansive lesson that involved exploring calculus concepts. The results show the learner constructed deep, profound, and novel notions of symmetry and curvature and emergent concepts of derivatives. In this paper, we report the way mathematical knowledge was understood and communicated through embodied ways of knowing and raise theoretical questions about the relationship between the body, the environment, and mathematical ideas.

INTRODUCTION

Mathematics educators often incorporate multiple representations into lessons to enhance learning. Recommendations for teaching algebra prioritize multiple visual representations to foster an understanding of relationships between graphs, pattern tables, and symbols (e.g. NCTM, 1991). However, not all bodies perceive all representations—students who are blind or have low vision (BLV) may not be able to perceive diagrams. Therefore, it is unsurprising that learning experiences for students with BLV are often inequitable and inadequate compared to sighted students (Stone & Brown, 2023). Mathematics education is starting to understand mathematical experience as embodied—mathematical ideas, physical objects, the environment, and bodies are entangled in mathematical knowing (e.g. de Freitas & Sinclair, 2014). Lakoff and Nuñez's (2000) work described mathematics as arising from bodily experience, while Alibali and Nathan (2012) described math ideas as communicated by or manifested in gestures. This case study examines a lesson conducted with a student with BLV that centered on a tactile representation in a university algebra class. We describe interactions between the student, instructor, mathematical ideas, and physical objects and connect our analysis to theories of embodied mathematics as they reach the classroom, envisioning an inclusive and expansive learning environment for all math students (Yu, Oslund, & Alshuli, 2024).

THEORETICAL FRAMEWORK

Math instruction has long centered visual images (NCTM, 1991), which may make math less accessible for BLV students (Sahin & Yorek, 2009). Students may also face teachers' lack of knowledge of accommodations, unawareness of Braille's challenges, or ineffective use of materials for teaching students with BLV (Bayram, et. al, 2015).

Language that assumes mathematics is visual perpetuates ableism in that it poses a barrier for instructional solutions for BLV learners. We have an ethical commitment to challenge ableist ideology and pedagogy strategically and practically by centering the learning of students with BLV (Tan, Padilla, & Lambert, 2022).

Historically, math was considered abstract and disembodied, but in recent years, mathematical knowing and learning has been theorized as embodied. An impetus for this has been the goal of democratizing mathematics learning and understanding ways bodies are societally positioned (Gerofsky, 2016; Sinclair & de Freitas, 2019). Research has studied the use of motion and other multi-sensory approaches to support students' learning of mathematics. For example, Gerofsky and Zebehazi (2022) used motion to help BLV elementary students notice mathematically important features of graphs. Developing an understanding of learning as embodied is important to our collective responsibility for equity. We analyzed data for this project to understand the entanglements between the student, instructor, mathematical concepts, and a set of wooden blocks. In the lesson, the student raised mathematical ideas that were surprisingly expansive, given the goals of the lesson. We asked, how does an understanding of math as embodied help us make sense of the mathematical concepts that emerged in the lesson?

METHODOLOGY

This study's context is a collegiate algebra course which focused heavily on visual representations. The student participant, Ian (pseudonym) was a male English language learner and international student from Central Asia in his first year as a computer science major. He was blind in one eye and had very limited vision in the other. Ian had no formal algebra or geometry coursework in his K-12 education. He was in a special section of this algebra course where he was taught one-on-one with the instructor, one of the researchers (Ball, 2000). Lessons were video recorded and data were analysed using iterative refinement cycles for video analysis (Lesh & Lehrer, 2000), including the transcription process, a review of the video data with the transcripts by at least two researchers together to identify emergent themes, and a review of the data by individual researchers to provide a deeper analysis of the identified themes. Portraiture was used to “create a narrative that bridges the realms of science and art, merging the systemic and careful description of good ethnography with the evocative resonance of fine literature” (Lawrence-Lightfoot, 2005).

RESULTS - THE FOCAL EPISODES

In this section we present three short episodes from an introductory lesson on parabolas and connect them to theories of embodied mathematics learning. Prior to this course the student, Ian, had no recollection of learning about parabolas and had just finished a unit on slope and linear equations (Oslund & Yu, 2023). This lesson used a puzzle consisting of a solid wood parabolic form and 15 wooden blocks that when stacked in the right sequence along the wooden form would result in a free-standing arch. The purpose of the lesson was to get Ian to feel the curvature and symmetry of a parabola.

However, the use of this novel puzzle allowed for a more expansive investigation on the properties of parabolas that extended well beyond the instructor's lesson goals.

Initial expected conversation

In the first episode, Ian began by removing the wooden blocks from the bag and placing them in three congruent groups saying, "I just like matching so I can get easier access with it." When asked if there was a pattern in his matching process Ian replied, "Not really," and began building the arch by taking two blocks from the groups and placing them in a symmetrical pattern along the wooden parabolic form (see Figure 1). While Ian did not recognize a specific pattern to his matching action, his building process reflected a form of embodied tactile symmetry as he grouped the blocks in pairs that he would grab two at a time. As Ian placed a block on one side of the arch, he would place a congruent block on the other side of the arch saying, "[This block] goes there, same pattern as well." When the instructor asked what was meant by same pattern Ian replied, "If you cut this through here [down the middle]," as he made a gesturing motion in which he used his right hand in a cutting motion down the axis of symmetry of the wooden parabolic form, "it's like, how do you say it? Symmetrical? Yes, symmetry when it's [symmetrical] both need to be same."

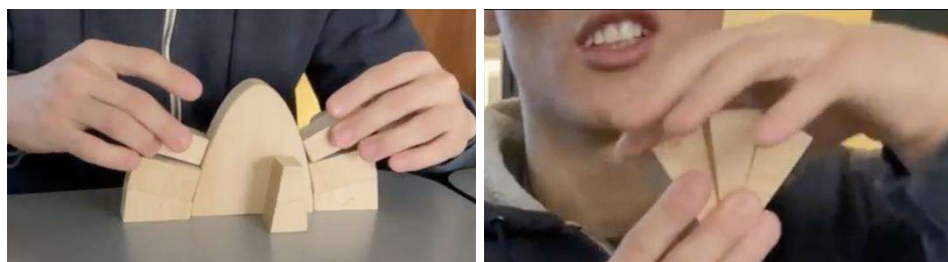


Figure 1: Symmetrical placement of blocks and gesture over the arch vertex

As he completed building the structure, Ian was asked if there was an order to stacking the blocks. He replied, "there is a specific pattern I would say, that the blocks have to go in." This specificity was indicated by how he felt the shapes of each block as he made the arch structure, often referring to physical characteristics of the blocks, like thickness and weight. He further explained that the three lightest blocks had to go on top and added, "this kind of slope makes sense as well, when [the keystone and adjacent blocks, or crown] are altogether on the top [of the arch], it makes sense... do you see [the three blocks'] angularity?" By angularity, Ian referred to how some blocks were more triangular and others more rectangular. As he explained, Ian used a gesture in which he cupped his fingers and moved them along the curve of three blocks (See Figure 1). Recognizing Ian's use of 'slope' and 'angularity' to describe the curve, the instructor introduced more formal terminology saying, "So the phrase we'll use is curvature. In particular, that part of that curve, which is called a parabola, that's the vertex. So, when you were talking about a very sharp curve, in mathematics they would say it has a high curvature."

This episode provides an example of the entanglement between the mathematical concepts, wooden puzzle, student, and instructor. The lesson was following the intended trajectory of the activity. Three foundational concepts of parabolas were discovered—symmetry, vertex, and curvature, mediated by his building ‘pattern.’ Ian displayed an embodied notion of symmetry as positional, based on the placement of congruent corresponding blocks as he built the arch. As Ian was touching, he was thinking (Alibali & Nathan, 2012). The instructor’s prompt to explain the ‘pattern’ sought to understand connections between Ian’s observed actions and emerging understandings. Ian articulated his embodied understanding of symmetry through partitioning as indicated by his gestural motion, leading to his use of more formal terminology like symmetrical and symmetry. Ian’s concept of curvature was also mediated by the shape of the blocks with more ‘angular’ blocks having high curvature as indicated by his gesturing motion, and more ‘heavy’ and ‘rectangular’ blocks corresponding to parts of the parabola with less curvature. The instructor provided the formal mathematical terminology reflected by Ian’s natural language.

Unplanned explorations of calculus

In the second episode, more complex notions of curvature were unexpectedly elicited when Ian was asked to further explain what he noticed and learned about the parabolic shape through the arch puzzle. Ian picked up the wooden arch form and traced his finger starting at the top, or vertex [point A], of the arch and ran his finger along the axis of symmetry to the bottom of the wooden form [Point M] saying, “Rise (See Figure 2).” Then he moved his finger along the bottom of the wooden arch to one of the base corners [point B] saying, “Run. If I divide [the rise] by [the run], I will get this straight line to connect point A to point B,” tracing his finger along the line that connected points A and B. As he traced his finger along the parabolic curve making up the edge of the wooden arch, he said, “But when I divide the rise by the run, I won’t get this curvy slant. How do you measure that curvy slope?”

Delightfully surprised the instructor said, “Oh, what a beautiful question.” Taking the wooden arch structure, he said, “I’m going to just hold [the blocks on the left side of the arch]. I’m gonna take apart [the right] half of [the wooden arch]. And I want you to take your finger and feel that inside curve (See Figure 2).” As Ian ran his finger along the arch’s inside curve the instructor said, “OK, now technically each of these blocks has a particular slope to it, right?” Ian responded affirmatively. The instructor continued, “So I’m going to call that a very localized slope.” As Ian placed his finger inside the bottom block the instructor said, “So your index finger right now, [on] that block right there, that straight curve has a slope. And that slope is different than this [top] block up here, right? That’s how you measure the curvature.” As Ian placed his finger on the inside vertical edge of the base block, the instructor asked, “Does that have a high slope or a low slope compared to [the top block] up here?”

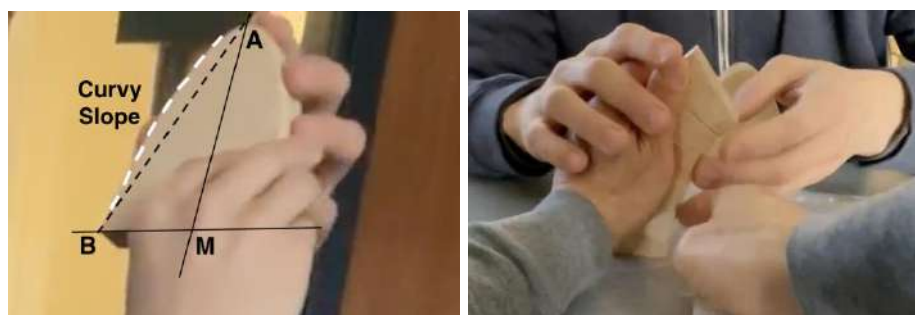


Figure 2: Constant vs curvy slope

As he tapped the bottom block in Figure 2 with his finger Ian said, “This [slope] is higher. The [lower block on the arch] is steeper.”

The instructor continued, “So the relationship between this high slope [on the lower block] and this slope [of the highest block] right here relates to the curvature. I would say the curvature is related to the change in slope.” As Ian affirmed his understanding the instructor continued, “when you get to calculus and I want you to remember this puzzle, right? That change in curvature is used with something called the second derivative... if we measured the rise and run [of the upper block], we could compare it to the rise and run of this block down here. The slopes. That's when we're talking about what's known as the first derivative in calculus.” As the instructor explained, gesturing, Ian would trace with his finger the vertical rise and horizontal run.

In this episode, an expanded understanding of a parabola emerged through entanglements between the student, instructor, math, and blocks, allowing for the development of calculus concepts that are introduced much later in the traditional course sequence. Ian used the solid parabola to wonder and compare the slope of the line from A to B with some unknown curvature measure while moving his fingers along the edge of the wooden form (see Figure 2), constructing the term, “curvy slope.” He recognized the inside edges of the arch’s blocks represent discrete slopes along the parabola. The instructor’s prompt to “feel” and compare the localized slopes of the arch’s lower and upper blocks led to more exploration of the relationship between slopes of lines along the parabola and curvature. The instructor used the slopes of the inside of the blocks as an initial metaphor of “an infinitely small interval” (Lakoff & Nuñez, 2000, p. 224.) describing how straight-line slopes of the “angular blocks” make up the “curvy slope” [parabolic curve]. The nature of the discourse shifted, becoming more formally mathematical, as the instructor connected Ian’s embodied experience and informal language with more formal terminology like “derivative in calculus.”

Extended explorations of Calculus

In the third episode, the lesson explored more explicit notions of the first derivative. Shifting his gestures to the wooden parabola, Ian traced his finger along the straight-line connecting vertex A and base corner B (See Figure 2) and said, “The first derivative.” He then ran his fingers along the “curvy slope” (outside edge of the form) and said, “second derivative, like around here.” Recognizing that Ian was confusing

the concept of instantaneous slope, or first derivative, with the line from A to B the instructor said, “So the first derivative, you're interested in the slope at a very localized level, very instant level, very small level.” The instructor referred to the wooden parabola and said, “If you hold this block up and put your fingers in two places along the side of the block. If you were to compare the slope [of the curve] of this thumb [Line M] versus the slope where your thumb is there [Line N], how would you compare those slopes? What do you think of?” See Figure 3.

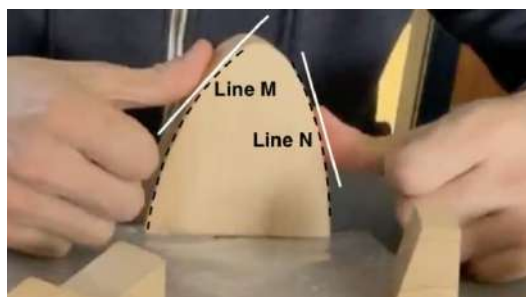


Figure 3: Exploring the derivatives on the wooden form

Tapping his left thumb at the point of tangency of line N Ian said, “This one's steeper.” Then tapping his right thumb at the point of tangency of line M Ian said, “this one is like. Less steep.” Referring to Ian’s thumbs on the wooden form the instructor clarified, “That idea of steepness that you're talking about where your left thumb [at Line N] is steeper and your right thumb [at Line M] is less steep at that point where your thumb is touching, that is all about the first derivative that you'll study in calculus. [It relates to] what we've been talking about [regarding] rate of change, like rise over run.”

Here there is a shift in the nature of the entanglements between student, instructor, mathematical concepts, and blocks. In the first two episodes the embodied connection between Ian and the blocks elicited math concepts described with informal language which was formalized by the instructor. In the third episode, in response to Ian’s partial understanding of derivatives, the instructor initiates an embodied connection between the block and Ian. The instructor let the ensuing conversation with more formal terminology focus on the connection between the wooden form, Ian’s thumbs, and the infinitesimal nature of the first derivative (Lakoff & Nuñez, 2000). The instructor’s repeated prompting for Ian to “feel” mathematical phenomena in all three episodes reflects an embodied pedagogy that extends beyond the typical use of physical manipulatives in mathematics instruction, more specifically, that mathematics could be felt if the object embodies the mathematics. The use of the wooden puzzle allowed Ian to feel and compare symmetry, steepness, curvature, and derivatives. However, Ian’s actions upon the wooden parabolic form and arch blocks also constituted meaningful mathematics interactions. In all three episodes, his grabbing, grouping, placing, touching, tapping, and tracing could be considered embodied mathematical actions that mediated increasingly complex mathematical concepts (Alibali & Nathan, 2012). Another embodied mathematical phenomenon in these episodes was the positional nature of mathematics. Symmetry was denoted by the placement of the

blocks in the wooden arch puzzle, curvature was explored by feeling the slope of the bottom and top blocks in the arch, and derivatives were explored thru the placement of Ian's thumbs on different parts of the parabolic form.

CONCLUSION

The student raised mathematical ideas that were surprisingly expansive, given the lesson's goals. We asked, how does understanding math as embodied help us make sense of the mathematical concepts that emerged? The video analysis helped us understand the entanglements between student, instructor, mathematical concepts, and the parabolic wooden block puzzle. Initially, the purpose of the lesson was to have the student feel the curvature and symmetry of a parabola through a learning experience involving a game-based tactile activity. However, as the student completely constructed the puzzle, this hands-on embodied experience allowed for a rich construction of mathematical concepts that expanded beyond the concept of parabola and into much-advanced mathematical concepts. As this interaction with the wooden arch puzzle was Ian's first formal educational exploration into quadratic parabolas, Ian brought ideas from calculus into being as he experienced the wooden blocks (Lakoff & Nuñez, 2000). These ideas were not told to him but emerged in the relationship between his embodied sensations and perception of the parabolic arch puzzle. The multimodal experience engaged Ian in a way that was not only accessible and inclusive, but also *expansive*. Because the instructor provided representations for algebra that expanded beyond the visual and symbolic, Ian was able to take the lead in an exploration of advanced math concepts. Ian's use of the informal language he invented to describe his perceptions intertwined with recently learned mathematical concepts (i.e. curvy slope) was supported and extended by the instructor, who embedded and introduced corresponding mathematical terminology.

This study has implications for the mathematics classroom. We posit that multimodal experiences can engage students in equitable ways, providing access and involvement to construct their mathematical understanding. It also supports the possibility of a lesson that is expansive, where students have the resources and authority and take the lead in their learning. Allowing students to invent and play with language for their learning is important to their learning. In expansive learning environments, this learning need not follow a traditional hierarchical order of math topics; students may take us beyond our assumptions into advanced and unexpected places.

References

- Alibali, M.W. & Nathan, M.J. (2012). Embodiment in mathematics teaching and learning: Evidence from learners' and teachers' gestures. *Journal of the Learning Sciences* 21(2), 247-286. <https://doi.org/10.1080/10508406.2011.611446>
- Ball, D. L. (2000). Working on the inside: Using one's own practice as a site for studying teaching and learning. In A. E. Kelly & R. A. Lesh (Eds.), *Handbook of research design in mathematics and science education*. Mahwah, NM: Lawrence Erlbaum Associates.

- Bayram, G.I., Corlu, M.S., Aydın, E., Ortaçtepe, E., Alapala, B. (2015). An exploratory study of visually impaired students' perceptions of inclusive mathematics education *British Journal of Visual Impairment* Vol. 33(3).
- de Freitas, E. & Sinclair, N. (2014). *Mathematics and the body: Material entanglements in the classroom*. NYC: Cambridge.
- Gerofsky, S. (2016). Approaches to embodied learning in mathematics. *Handbook of international research in mathematics education*, 3rd edition. New York: Routledge.
- Gerofsky, S. & Zebehazy, K.T. (2022). Enhancing mathematical noticing of graphs through movement, voice, and metaphor: An intervention with two students with visual impairment. *British Journal of Visual Impairment*, 40(2), 209-221. <http://doi.org/10.1177/0264619620963516>
- Lakoff, G. & Núñez, R.E. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being*. NYC: Basic books.
- Lawrence-Lightfoot, S. (2005). Reflections on Portraiture: A Dialogue Between Art and Science. *Qualitative Inquiry*, 11(1), 3-15. <https://doi.org/10.1177/1077800404270955>
- Lesh, R., & Lehrer, R. (2000). Iterative refinement cycles for videotape analysis of conceptual change. In A. E. Kelly & R. A. Lesh (Eds.), *Handbook of research design in mathematics and science education*. Mahwah, NJ: Lawrence Erlbaum Associates
- NCTM (1991). *Professional standards for teaching mathematics*. Reston, VA: NCTM.
- Oslund, J. A. & Yu, P. W., (2023). Understanding slope thru tactile and whole-body experiences. In *Proceedings of the forty-fifth annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education (Vol 2)*, 675-679. University of Nevada, Reno.
- Sahin, M., & Yorek, N. (2009). Teaching science to visually impaired students: A small-scale qualitative study. *US-China Education Review*, 6(4), 19–26.
- Sinclair, N. & de Freitas, E. (2019). Body studies in mathematics education: Diverse scales of mattering. *ZDM* 51, 227-237. <https://doi.org/10.1007/s11858-019-01052-w>
- Stone, W. B., & Brown, D. (2023). Anyone can learn universal design: An interdisciplinary course centered around blindness and visual impairment. *Journal of Postsecondary Education and Disability*, 2023, 36(1), 65-74.
- Tan, P., Padilla, A., & Lambert, R. (2022). A Critical Review of Educator and Disability Research in Mathematics Education: A Decade of Dehumanizing Waves and Humanizing Wakes. *Review of Educational Research*, 92(6), 871-910.
- Yu, P. W., Oslund, J. A. & Alshuli, T.A. (2024). Understanding function transformations through tactile representations and gesturing with visually impaired and sighted students. In *Proceedings of the forty-sixth annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Kent State University.

AXES WITHOUT NUMBERS: WHAT FUNCTION CAN THIS BE?

Rina Zazkis

Simon Fraser University

A line is drawn in a coordinate system, with two perpendicular axes labeled X and Y, but the axes are not labelled with numbers. What function corresponds to this line? This paper reports on how prospective teachers envision guiding students through this question. The data consists of participants' responses to a scripting task which required them to present their ideas for instruction in the form of a dialogue between a teacher and student-character. The results demonstrate that while most participants focused on student thinking, only a few acknowledged a possibility of non-homogeneous coordinates.

INTRODUCTION

Our study is inspired by an exchange which occurred in a secondary mathematics classroom: The teacher sketched a graph, similar to the one shown in Figure 1, and asked students, what function the graph represented. Initially, students claimed that the task was impossible, as no numbers were given. The teacher then asked students to put numbers on the axes and the class agreed that the graph “looked like” $y = 2x$. We wanted prospective teachers to consider different approaches to the above scenario. So, we created a scripting task in which prospective teachers participating in our study were asked to address the students' claim by composing a script for an imagined discussion.

ON LINEAR FUNCTIONS

Our investigation centers on an exploration of linear functions. The conventional curricular approach involves students learning to graph linear functions given a slope and Y-intercept, or equation of the form $y = mx + b$, and to find the corresponding linear function given two points on a line. This mathematical content is considered in the Cartesian system of coordinates. That is, the underlying assumption for exploring linear functions is that their graphs are constructed in a homogeneous coordinate system – in a system with perpendicular axes and homogeneous units on both axes.

Research on students' understandings of linear functions focused on what Moschkovich et al. (1993) called the “Cartesian connection.” That is, connecting the set of points on a line to the set of solutions of the corresponding linear equation. This research outlined students' difficulty in moving between different representations of linear functions and using available information provided by each representation (e.g., Davis, 2007; Knuth, 2000). Researchers also pointed to students' tendency to visually estimate the value of points on a graph, even when the information for finding precise values was available (Glen & Zazkis, 2021).

However, graphing software, such as Desmos, allows for flexibility that is not usually present with paper and pencil graphing. Acknowledging this flexibility, Zaslavsky et al. (2002) presented participants with a task in which they were asked to determine the slope of the function $f(x) = x$ graphed in a non-homogeneous system of coordinates (the unit on the X-axis was 3 times longer than the unit on the Y-axis). Zaslavsky et al. (2002) indicated that many participants relied on their image of $f(x) = x$ forming a 45-degree angle with the X-axis and thus had trouble with a task scenario where this expected image was not consistent with the task. Here we explore a related phenomenon; whether teachers can flexibly attend to the size of units in a teaching scenario they were asked to imagine.

THEORETICAL FRAMING: THE “MOST” FRAMEWORK

Leatham et al. (2015) introduced the theoretical construct: Mathematically significant pedagogical Opportunities to build on Student Thinking (MOSTs). This framework facilitates teachers’ ability to identify “when it is productive to act on student thinking at the moment in which it occurs” (p. 90); student thinking that provides such a productive moment is called a MOST. According to Leatham et al. (2015), a MOST is a student contribution that is characterized by three attributes: (a) it reveals the *student’s mathematical thinking*, (b) it is related to a concept that is *mathematically significant*, and (c) it provides a *pedagogical opportunity*. There are three supporting criteria for deciding whether a particular contribution is a MOST. For attribute (a) it is teacher’s ability to articulate student thinking, as related to the mathematics of the lesson; for attribute (b), the mathematical issue should be related to the topic and accessible to students; and attribute (c) requires the teacher to decide whether responding to the student accords with the learning goals for the lesson.

Since its inception, the MOST framework has informed a number of studies in mathematics education. One common application of the framework involves gaging the growth of prospective teachers over a course (Fernández et al., 2020) or following a specific intervention (Stockero et al., 2017; Stockero, 2021). In these studies, the accuracy with which prospective teachers can identify MOSTs is used to study the effectiveness of certain activities in teacher education programs, for example, writing and receiving feedback on narrative teaching episodes (Fernández et al., 2020); classroom observations with an emphasis on “important instances”, followed by consultations with experienced instructors (Stockero et al., 2017); and explicit instruction on the MOST framework (Stockero, 2021). In some cases, specific pedagogical interventions positively affected teachers’ ability to identify MOSTs (Stockero, 2021).

Leatham et al. (2015) noted that pedagogical opportunities can arise from many types of student thinking. These include: “(a) a correct answer with novel reasoning, (b) an incorrect answer that involves a common or mathematically rich misconception, (c) a mathematical contradiction, (d) incomplete or incorrect reasoning, and (e) *why* or generalizing questions” (p. 100, italics in original). The student thinking featured in

this study related to categories (d) and (e): it features a student claim that the task is impossible and a reasonable but imprecise answer from another student. In exploring teachers' reactions to student claims, we address the following research question: *How do teachers leverage the pedagogical opportunity afforded by students claims in the case of a linear function graphed on unlabeled axes?*

METHODS

Scripting tasks

Scripting tasks were first described in Zazkis et al. (2009, 2013) as a tool for research and professional development. Briefly, participants write a dialogue that describes the way in which they envision interaction between a teacher and students. A task typically begins with a brief initial conversation that is closely related to teaching and learning mathematics. This initial dialogue, along with details of the surrounding classroom context, is called a prompt. A prompt can take many forms, such as a student error, a question, or an argument among students. Scriptwriters complete the task by continuing the given dialogue between the teacher- and student-character(s), and in doing so, address the mathematical issue introduced in the prompt.

Scriptwriting was used as a methodological tool in a variety of studies in mathematics education (e.g., Kercher et al., 2023; Koichu & Zazkis, 2013; Marmur et al., 2020). Researchers argued that scripts composed by prospective teachers serve as a valuable tool as a preparation for practice, and provide a lens for researchers to examine participants mathematical knowledge and pedagogical inclinations.

We identified student responses introduced in the prompt to the scripting task as MOSTs, as they satisfied the three criteria described above. We agree with Son and Crespo (2009) that responding to a teaching-scenario task might simulate “how mathematical work arises in the context of teaching” (p. 243). Engaging in scriptwriting presents a way for prospective teachers to examine their pedagogical approaches. Further, we note that providing teachers with time to respond to a teaching scenario in the form of a written dialogue provides valuable preparation for the “thinking on your feet” that is required in an actual classroom situation.

The Task

The beginning of the dialogue presented in Figure 1 served as a prompt for a scripting task. In addition to completing the dialogue (Part 1 of the task), the participants were asked to describe their pedagogical choices (Part 2) and also note if their personal understanding of the mathematical content differed from what they chose to discuss with students (optional Part 3 of the task).

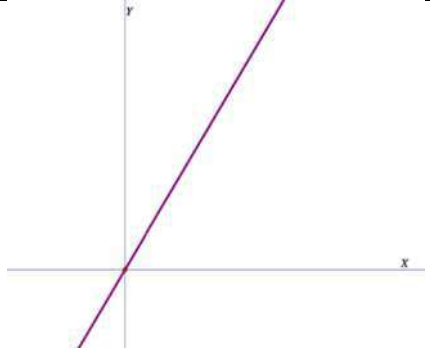
<p>Your students are considering the following graph.</p> <p>Teacher: What is the function that corresponds to this graph?</p> <p>Alex: There is no way of knowing this, as there are no numbers.</p> <p>Teacher: Indeed, the numbers are hidden, but we still can draw some conclusions.</p> <p>Jamie: It looks like $y = 2x$</p> <p>Teacher: ...</p>	
---	--

Figure 1: Prompt for the Scripting task

Participants

Participants in this study were 14 prospective secondary school teachers (referred to here as P1 through P14). At the time of data collection, they were enrolled in a “methods” course in their final term of a teaching certification program. The course focused on the didactical structure of mathematical topics that are central in the secondary school curriculum. However, the course also focused on enhancing the participants’ understanding of mathematics by engaging them in problem solving and connecting ideas of tertiary mathematics to secondary school teaching.

The participants completed several scripting tasks as part of their regular coursework. Their responses to one such task, along with their responses to the accompanying discussion prompts, serve as the dataset for this report.

Data Analysis

We read the scripts (Part 1 of the task) several times and coded the teacher’s responses. We first noted whether the exchange was identified by the scriptwriter as a MOST, that is, whether the teacher considered the presented student idea as an opportunity for valuable mathematical extension. When this happened, we focused on the strategies chosen by participants. We further attended to their commentary in parts 2 and 3 as a potential demonstration of their personal mathematical knowledge not visible in the dialogue, and their justifications for their chosen pedagogical approaches.

FINDINGS

The value of the y-intercept was clearly and correctly determined by all the student-characters in all the scripts, and it is implied, (though not explicitly) in Jamie’s suggestion that the function “looks like $y = 2x$ ”. As such, we directed our analysis to the discussion of slope as it appeared in the scripts.

Out of 14 participants, three did not consider the exchange presented in the prompt as a MOST. Rather than attending to the features of the graph available in the sketch, and the estimation of the value of the slope, these scriptwriters focused their instructional

approach on revisiting previously learned content and determining the precise equation that described the graph. For example, in P11's script, the student-characters initially attempted to find the x -value that corresponds to $y = 2$. Then the x -value was estimated as 1.1 and the slope was determined with the aid of a calculator (Figure 2a). However, following the teacher's request to avoid the use of a calculator, students determined what value of y corresponds to $x = 1$, and concluded, by measurement, that the sought function was $y = 1.8x$ (See Figure 2b). This was further confirmed by another pair of student-characters, who attended to $x = 2$ and measured the corresponding y value as 3.6.

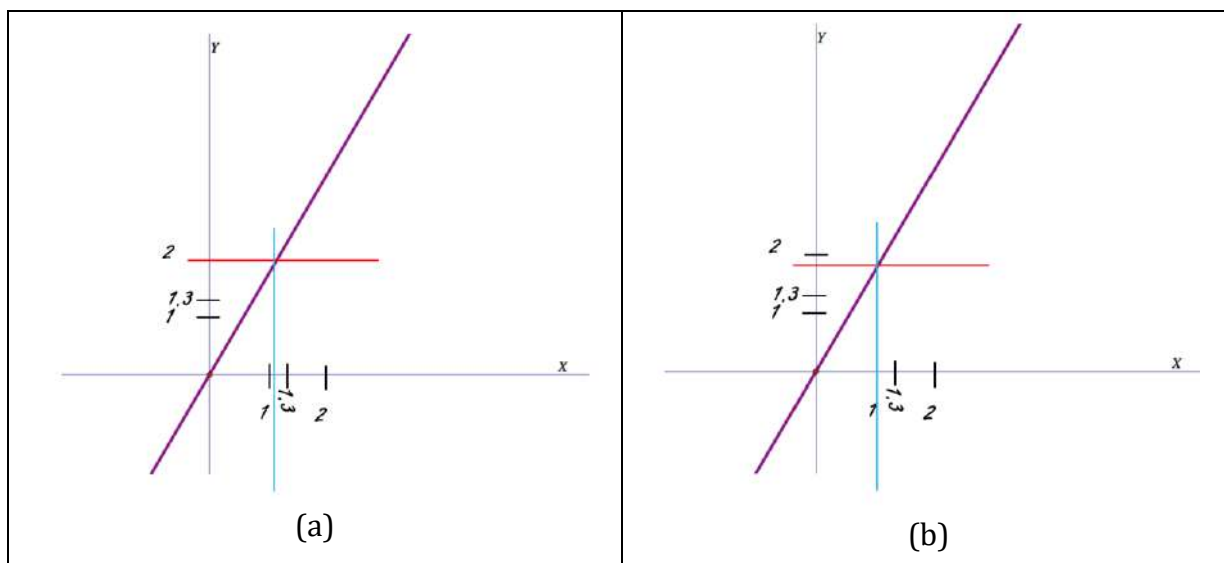


Figure 2: Finding the exact value of the slope

The following excerpt attends to this approach and extends it, confirming the result.

- Teacher: Alex, can you think of a way to make an estimation of the slope without using your calculator.
- Jamie: I think I have an idea! Move the line down!
- Alex: Move the line down where, I'm so confused.
- Jamie: Well since the slope is equal to the rise of the line divided by the run of the line... if we adjust the red line so that the blue line is exactly at $x = 1$, then we will be dividing by 1, and thus our slope will be where our red line intersects the y -axis. Let me adjust the lines and show you.
- Jamie: And [...] as long as we draw the line that intercepts the y -axis, red in our case, and the line that intercepts the x -axis, blue in our case, carefully, we could have called our increments anything we wanted and it would have resulted in the same result for the slope of our line!
- Alex: Right that makes sense! I guess if we would have called our first increment 5 cm, and the second increments 10 cm, we would get a slope of about $\frac{9}{5}$ and if I'm not mistaken that's $1\frac{4}{5}$ or 1.8!

Teacher: This has all been very great work.

Of the participants who attended to the MOST, 7 out of 14 focused on the students' explanation of $y = 2x$, having concluded that the y-intercept was zero. These scriptwriters had their student-characters examine various graphs of linear functions to confirm, by eyeballing, that the slope was between 1.5 and 3, or that the slope was positive and greater than 1. The latter conclusion explicitly attends to the fact that the line $y = x$ bisects the angle between the axes and that the line given in the task is steeper. In these discussions we find valuable pedagogical opportunities for supporting students' ability to estimate slope based on visual examinations of graphs and to justify an estimated value. Regardless, the pedagogical opportunity can be extended further.

P13 alludes to the possibility of flexible assignment of coordinates by acknowledging the possibility of different scales on the axes:

Teacher: If we assume that the scale on each of the axes is the same, then your initial guess of $y=2x$ is reasonable because the y-intercept is zero and the slope is a positive number. Great work!

However, the above appeared on the last lines of the script and no alternatives for the acknowledged assumption were considered in the script and in the commentary.

Five scripts explicitly attended to the possibility of non-homogenous coordinate systems, that is, the unit on the X-axis is of different length than the unit on the Y-axis. This approach is exemplified in the following two excerpts.

P4's script asked two student-characters, Alex and Jamie, to label the X and Y axes, respectively. They each chose different units.

T: Well both of you marked your own axis separate from each other so they were very likely not the same. [...] This is called the *scale*. Alex, can you tell me what was the relationship of the x-axis scale and y-axis scale of the first three markings you made?

A: They were all the same...

T: Exactly! And when you two made the scale separately the scale was different. We assume a lot of the time that the scale has to be the same on both axes, but as we have seen that is not true.

J: So the answer is not $y=2x$?

T: Not quite. But we can still say something about the graph.

In P2's script, the student characters initially claimed that the slope must be greater than 1. Then, the teacher-character challenged them to consider the scale on the axes.

Teacher: Perhaps we could think about it this way – what is the scale on the x and y-axis?

Alex: ... There is no scale. Like I said earlier, there are no numbers on the axes.

- Teacher: Precisely! So if we have no idea what the scale of the x and y-axis are, how do we know what the slope of the line is?
- Jamie: Oh, I see! Since we don't know the scale, it could be anything. The scales on the x and y-axis don't even have to be the same, right? Like the x-axis could go up by 2's whereas the y-axis could go up by 1's. [...]
- Alex: Hmm... So technically the slope of the line could be anything depending on how we adjusted the scale of the x and y-axis.
- Teacher: Are you sure it could be anything?
- Alex: Ok, well not ANYTHING, anything that is positive I guess. It has an increasing slope, so it can't be negative.

In her comments, P2 wrote:

Understanding that data can be manipulated in many ways is a powerful lesson for students to learn. Being able to recognize the way in which statisticians, mathematicians and charlatans can manipulate graphs to better convey their particular point will help them better critically analyse data in the future.

We appreciate this comment in support of students' mathematical education beyond the particular task at hand.

DISCUSSION, CONTRIBUTIONS AND CONCLUSIONS

Prior research acknowledged that challenging basic assumptions is a valuable mathematical engagement (Zazkis, 2008). In particular, the decimal representation of numbers and the Cartesian coordinate system were mentioned among the conventions that can be changed and challenged. This resulted in increasing understanding of related assumptions. For example, the distance formula is not applicable when the coordinate system is not homogeneous. In fact, the notion of distance is meaningless when working with non-homogeneous coordinates.

In our study we addressed the following research question: *How do teachers leverage the pedagogical opportunity afforded by students' claims in the case of a linear function graphed on unlabeled axes?* We found that most participants utilize the presented pedagogical opportunity and support, in their envisioned teaching approaches, students' connections between the algebraic and graphical representations. However, only a few approached the task flexibly by rejecting the conventional restriction imposed by Cartesian coordinates. This task served as a springboard for the follow-up lesson where participants worked with Affine coordinates; that is, not only considering different scales on the axes but also breaking the convention of axes being perpendicular.

In conclusion, contributing to mathematics teacher education, the scripting task discussed in this report provided a valuable avenue to examine, and then to extend, participants' personal knowledge, both mathematical and pedagogical, and enhanced the ways of bringing this knowledge to a classroom.

References

- Davis, J. (2007). Real-world contexts, multiple representations, student-invented terminology, and y-intercept. *Mathematical Thinking and Learning*, 9, 387–418.
- Fernández, C., Llinares, S., & Rojas, Y. (2020). Prospective mathematics teachers' development of noticing in an online teacher education program. *ZDM Mathematics Education*, 52, 959-972.
- Glen, L., & Zazkis, R. (2021). On linear functions and their graphs: Refining the Cartesian connection. *International Journal of Science and Mathematics Education*, 19(7), 1485-1504.
- Kercher, A., Bergman, A. & Zazkis, R. (2023). What is geometric about geometric sequences? Informal justifications for mathematical terminology. *Canadian Journal of Science, Mathematics and Technology Education*, 23, 48-65
- Koichu, B. & Zazkis, R. (2013). Decoding a proof of Fermat's Little Theorem via script writing. *Journal of Mathematical Behavior*, 32 (364-376).
- Knuth, E. J. (2000). Student understanding of the Cartesian Connection: An exploratory study. *Journal for Research in Mathematics Education*, 31(4), 500–507.
- Leatham, K. R., Peterson, B. E., Stockero, S. L., & Van Zoest, L. R. (2015). Conceptualizing mathematically significant pedagogical opportunities to build on student thinking. *Journal for Research in Mathematics Education*, 46(1), 88-124.
- Marmur, O., Yan, K., & Zazkis, R. (2020). Fraction images: The case of six and a half. *Research in Mathematics Education*, 22(1), 22-47.
- Moschkovich, J., Schoenfeld, A., & Arcavi, A. (1993). Aspects of understanding: On multiple perspectives and representations of linear relations and connections among them. In T. A. Romberg, E. Fennema, & T. P. Carpenter (Eds.), *Integrating research on the graphical representations of functions* (pp. 69–100). Hillside: Erlbaum.
- Son, J. W., & Crespo, S. (2009). Prospective teachers' reasoning and response to a student's non-traditional strategy when dividing fractions. *Journal of Mathematics Teacher Education*, 12(4), 235-261.
- Stockero, S. L. (2021). Transferability of teacher noticing. *ZDM Mathematics Education*, 53(1), 73-84.
- Stockero, S. L., Rupnow, R. L., & Pascoe, A. E. (2017). Learning to notice important student mathematical thinking in complex classroom interactions. *Teaching and Teacher Education*, 63, 384-395.
- Zazkis, R. (2008). Examples as tools in mathematics teacher education. In D. Tirosh (Ed.) *Tools in Mathematics Teacher Education*. (in Handbook for Mathematics Teacher Education, Vol. 2, pp. 135-156). Rotterdam, Netherlands: Sense Publishers.
- Zazkis, R., Liljedahl, P. & Sinclair, N. (2009). Lesson Plays: Planning teaching vs. teaching planning. *For the Learning of Mathematics*, 29(1), 40-47.
- Zazkis, R., Sinclair, N., & Liljedahl, P. (2013). *Lesson Play in Mathematics Education: A tool for research and professional development*. Dordrecht, Netherlands: Springer.

URBANIZATION AS A CATALYST OF FRACTION UNDERSTANDING: A CASE STUDY

Sian Zelbo

Teachers College, Columbia University

This historical case study illustrates how demographic and economic changes in early twentieth-century America required ordinary people to engage with fractions in new and more sophisticated ways. Urbanization and industrialization took more and more people out of the routine contexts of pre-industrial commercial life and demanded that they engage with complex industrial processes and financial systems. Through an analysis of word problems in annual standardized high school mathematics exams and using a framework based on the work of Siegler and others, preliminary findings suggest that these new contexts pushed ordinary people to conceptualize fractions in more complex ways.

INTRODUCTION

The demographic and economic transformations of early twentieth-century America created new demands in how ordinary people engaged with mathematical concepts. As people moved from rural areas and small towns into cities, and as immigrants poured into the country seeking factory work, they encountered standardized systems of mass production that required more precise measurements and more complex record-keeping. This study examines the transformation in societal mathematical demands in this period through an analysis of annual standardized high school exams. Drawing from a historical record of New York state exams spanning nearly 160 years, initial analysis suggests that the shift to an urban industrial economy pushed ordinary people to conceptualize fractions more consistently as abstract magnitudes rather than as countable parts of a whole, and toward an understanding of fractional magnitudes as points on a continuous number line.

This historical case study also illustrates a broader principle. While mathematical understanding evolves differently across cultures and time periods, examining specific historical contexts can reveal patterns that have practical implications for modern education. This contextual perspective helps explain, for example, why certain concepts consistently challenge learners. Indeed, as mathematics historian and educator David Eugene Smith observed about fraction division, “it took the world thousands of years to learn this process” (Smith, 1909, p. 67-68)—suggesting that certain complex mathematical ideas require sustained engagement to take root, a process that education must respect rather than attempt to bypass. The conceptualization of fractions as abstract magnitudes and as points on a continuous number line represents another such historically challenging transformation—one that

this study suggests may be shaped not merely through improved mathematics instruction but through sustained engagement with new cultural demands.

Building on this historical context, this study has two primary research goals: (1) To analyze how changes in standardized mathematics assessments between 1890-1930 reflected new mathematical demands as New York's population urbanized; and (2) to provide an example of how new cultural contexts and practical demands can influence how people conceptualize mathematical ideas.

THEORETICAL FRAMEWORK

Research on mathematical cognition has documented how, with educational interventions, individuals can develop an increasingly sophisticated understanding of numbers, particularly in their conceptualization of fractions. Two fundamental conceptual transformations are particularly relevant to this study: the progression from viewing fractions as parts of wholes to understanding them as abstract magnitudes that can be ordered and compared, and the evolution from discrete to continuous thinking about the real number line. Studies consistently show that both children and adults exhibit a persistent whole number bias that makes these conceptual shifts difficult (Siegler et al., 2013; Ni & Zhou, 2005). Learners have strong intuitions about numbers that mirror whole-number properties, such as that numbers are evenly spaced, that they can be counted one by one, that each number has a “next” number, and that each number has a unique symbolic representation (Ni & Zhou, 2005). These deeply held intuitions, while useful for early arithmetic, present a fundamental challenge to developing a more sophisticated understanding of fractions (Siegler et al., 2013; Ni & Zhou, 2005).

A critical conceptual shift is that students must learn to view rational numbers in the form a/b not just as a relationship between two whole numbers or as countable parts of a single, divided whole. Instead they must conceptualize $7/8$ or $6/5$, for example, as abstract magnitudes that can be ordered and placed on a number line (Siegler et al., 2011). A student exhibits the more limited understanding of fractions, for example, when he believes that the fraction $4/3$ has no meaning because you cannot take four parts of an object that is divided into three parts (Mack, 1993; see also, for example, Ahl and Helenius, 2024 with an example of how instructional approaches can facilitate or inhibit the shift). This shift to thinking of fractions as magnitudes is difficult because it requires a fundamental reorganization of one's concept of number. The cognitive progression to an ordered mental number line of rational numbers is slow, complex, and often incomplete (e.g., Schneider & Siegler, 2010).

A second aspect of students' conceptual development of fractions is understanding that infinitely many numbers exist between any two points and the related idea that the real number line is continuous (Vamvakoussi & Vosniadou, 2004; Vamvakoussi & Vosniadou, 2010). The understanding of fractions as magnitudes is necessary but not sufficient for students to make this second conceptual leap (McMullen et al., 2015). The development of this understanding, too, is a slow and complex process that occurs

over years of mathematical study, as González-Forte et al. (2021) showed, if it happens at all (Stigler, Givvin, & Thompson, 2010).

The challenge of moving from discrete to continuous thinking has been theorized in a particularly useful way by Castillo-Garsow et al. (2013). They identified two patterns in how students conceptualize change: “chunky” thinking, where change occurs in discrete steps (like filling a bucket cup by cup), and “smooth” thinking, where change is continuous (like filling a bottle from a hose). This distinction is evident in early number line experiences, where students typically encounter discrete whole-number units and rulers marked at regular, countable intervals. Students may not even be confronted with the density property of the real numbers until they begin to work with repeating decimals and irrational numbers. Yet, the limited calculation experiences students have with such numbers lead them, and indeed many adults as well, to have underdeveloped intuitions about the real number system (Sirotic & Zazkis, 2007).

This framework about the complex process of fraction concept development provides a valuable lens for examining how demographic and economic changes demanded more sophisticated forms of mathematical understanding. Just as learners require sustained engagement with new mathematical experiences to develop a more complete understanding of fractions, ordinary people of the early twentieth century may have developed more complex ways of thinking about fractions as they engaged with the new mathematical demands of urban industrial life.

RESEARCH METHODS

This study examines societal mathematical demands of a specific time and place through an analysis of the New York state examination system. Drawing from the Jefferson Math Projects’s comprehensive digital archives of New York State Regents exams dating back to 1866, initial analysis has identified a significant transformation in the complexity of questions that took place in the first decades of the twentieth century. The current phase of research focuses on exams from the years 1890 to 1930, a period that precedes and includes this critical shift in mathematical complexity, with particular attention to fraction conceptualization.

The early twentieth century marked an unprecedented democratization of secondary education in New York, with high school enrollment surging from 100,000 students in 1900 to 350,000 by 1925 (State Education Department, 1965). As the state exams evolved from selective evaluation tools to broader assessments of students’ mathematical skills, they provide a particularly valuable lens for understanding the mathematical demands placed on ordinary high school students.

For this study, I examined the June administrations of the state’s lowest-level high school mathematics exams—Arithmetic and Elementary Algebra—because these exams reached the broadest possible student population. The ongoing analysis of the available exams in the period selected proceeds along two dimensions:

Cultural/Technological Context: The study documented the types of measuring tools implied by the problems, any units of measurement in word problems, the level of precision required in measurements, and any references to emerging commercial contexts or technologies.

Mathematical Conceptualization: The study identified problems that reflect part/whole versus magnitude understanding of fractions, noted the progression from discrete, countable units in familiar contexts to more abstract representations of fractional amounts, and noted when the context of a word problem suggests mathematical engagement with a continuous process.

Through this systematic examination, patterns have emerged that suggest connections between new cultural contexts and shifts in mathematical demands for conceptualizing fractions. Future phases of research will examine additional primary sources to help paint a more complete picture of this educational transformation.

ANALYSIS

Pre-Industrial Era: Discrete Thinking in Simple, Routine Contexts

In 1900, New York straddled old and new economies. While over thirty percent of the state's seven million residents lived on farms and in small towns (U.S. Census Bureau, 1901, p. lxv), the population was rapidly urbanizing. New York City's population, which had grown modestly in the 1880s, more than doubled in the 1890s from 1.5 to 3.4 million residents (U.S. Census Bureau, 1901, p. lxix), with many immigrants arriving and with people moving into cities for wage work.

Pre-1900 arithmetic exams reflected a world where mathematical contexts were routine and compartmentalized, each with its customary units and notations. Word problems from this era often conceptualized fractions as part/whole relationships rather than as abstract magnitudes. For example, an 1892 Arithmetic exam describes a "railway train [that] runs $\frac{2}{3}$ of a mile in $\frac{4}{5}$ of a minute." The phrasing "of a mile" and "of a minute" linguistically reinforces these fractions as parts of divided wholes. Even when fractions were treated as magnitudes, they often remained tied to tangible, countable parts of a whole—like yards of fabric measured to the nearest quarter-yard or fencing measured to the nearest half-rod.

Before 1900, state exams largely avoided decimal notation, instead expressing sometimes even dollar amounts as fractions or mixed numbers. The 1890 Arithmetic exam, for example, included a problem about a carpet purchased for "\$26 $\frac{1}{4}$ " and sold "at a gain of $\frac{1}{4}$ on each yard." Dollar amounts in these early exams that were expressed to the half-cent would mix notation, such as an 1891 problem that involved the payment of "\$648.12 $\frac{1}{2}$ in taxes." This mixed notation suggests that the decimal part of a dollar was viewed as a whole number of cents rather than as hundredths of a dollar. The first use of decimal notation outside of the context of money among the Arithmetic exams in this study involved metric units (an 1894 Arithmetic exam question involving layers of bricks that average "6.5 centimeters in thickness"). When

decimal notation was used in the 1890s, it was largely confined to just a few contexts, including money, metric units, and occasional square root calculations. Because of the focus on routine calculations and the general tendency to minimize and compartmentalize decimal notation, these early Arithmetic exam problems rarely required students to convert, order, compare, or estimate fractional amounts—mathematical operations that might have fostered a more sophisticated understanding of fractions as magnitudes on a number line.

Elementary Algebra exams of the earlier era, despite being designed for college-bound students, also gave students few opportunities to engage with fractions in more complex ways. The 1890 exam exemplifies this approach. While it included rational expressions and radicals, students mainly performed algebraic manipulations—such as simplifying complicated expressions—without the need to consider the value of the expressions. (One exception to the tendency was an 1891 question that asked students which was larger, $\sqrt[3]{3}$ or $\sqrt[4]{4}$.) With the focus on algebraic manipulation, only one of the thirteen problems on the 1890 exam involved “real-world” application: a room measurement problem that produced a quadratic equation with whole number solutions. The first algebra word problem in this study with common fractions appeared in 1895, and fractions were used in a conceptually simple way, emphasizing the part/whole relationship: “The sum of two numbers is 20 and one half the larger is equal to three fourths the smaller” Like the Arithmetic exams of the time, these Elementary Algebra tests reflected limited opportunities for students to consider fractions as magnitudes. In fact, in some ways they lagged behind the Arithmetic exams of the same era, perhaps because the algebra curriculum was not as directly tied to immediate, real-world needs.

The Industrial Era: Higher Demands in the Conceptualization of Fractions

Industrial manufacturing and mass production transformed not just how Americans worked, but how they calculated. The precise measurements and complex calculations demanded by factories and modern finance pushed workers—and state exams—toward deeper mathematical understanding, especially of fractions. While these changes in exam content likely reflected multiple influences—including evolving pedagogical theories and educational reforms—industrialization presented genuinely novel mathematical demands involving contexts that ordinary people in the pre-industrial era had never encountered in their daily lives.

One aspect of the transformation in exams was that word problems shifted from focusing exclusively on small-scale commerce to include references to large-scale commercial operations. Examples included a 1915 Arithmetic problem, which had students calculate percentage savings for shipping freight through the newly opened Panama Canal versus by railroad, requiring precise decimal work in a context of international commerce. A 1921 Arithmetic exam had students analyze corporate payroll expenditures through sequential percentage changes. A 1925 Arithmetic problem required calculating automobile depreciation over three years—a calculation

necessary for the federal income tax, which had been newly instituted in 1913. These changes are not necessarily mathematical but show that exam writers were responsive to the cultural context of a more complex society.

The transformation also manifested in the increasing and more flexible use of decimals. This shift is significant because decimal numbers are easier to order and compare, and the flexible use of fractional numbers in different forms, therefore, is key to developing a more sophisticated understanding of fractions as magnitudes. The 1901 Arithmetic exam, for example, included a problem in which the perimeter of a circular park was given as 314.16 rods, and students needed to find the area. The use of decimals outside of the context of money and metric units in this problem was novel and signaled a slow move toward the broader use of decimals. By 1910, for the first time, an Arithmetic exam included questions asking students explicitly to convert abstract values between fractions, decimals, and percentage forms. While the progress in the use of decimals was not consistent year to year, the difference between the exams from the 1890s and the 1930s is striking. The 1890 exam included only five problems with fractional numbers out of thirteen total problems. The five problems all included simple denominators in conceptually simple and routine contexts, and the only use of decimals was in the context of money. In the 1930 exam, by contrast, 29 of the 33 questions involved fractions, decimals, and percents in a variety of contexts, often multi-stepped and with mixed representations within a single problem.

Exams from this transitional period also hinted at the continuous nature of the number line for the first time. A 1916 Arithmetic exam problem, for example, tracked a warship's speed over four hours using decimal measurements (19.5, 21.75, 22.2, and 22.9 miles). While the ship's distance is tracked only once an hour and could be viewed as "chunky" not "smooth" thinking, the context of a ship varying its speed, and the decimal precision of the answer (21.5875 miles per hour), at least hint at a continuous number line. This contrasted with more traditional rate problems that used whole numbers and simple fractions for discrete tasks like workers digging trenches or plowing fields. Similarly, a 1917 Elementary Algebra exam problem about "a printing press [that] did a certain piece of work in 4 hours" also represents a subtle but significant shift toward continuous thinking. While a printing press produces discrete physical objects (like pages or newspapers), its operation represents a continuous process—the steady flow of paper and ink producing output at a constant rate. Unlike traditional rate problems involving discrete units of labor, this industrial context pushed students to think about production as a continuous flow over time rather than a sum of discrete units.

The 1918 Elementary Algebra exam pushed this idea of continuity still further by having students analyze, to five decimal places, the error introduced when approximating the rational expression $(1+x)/(1+y)$ with the simpler expression $1 + x - y$ for small values of x and y . The need to understand small variations and errors at this level of precision directly reflected the demands of industrial quality control—where

workers needed to understand the cumulative effects of small measurement errors in mass production processes. This type of mathematical thinking marked a significant departure from the simpler fractional calculations required in pre-industrial contexts.

Despite these advances in working with fractions as abstract quantities, even by 1930 the shift was incomplete. Text questions in this last decade of the study only began to hint at continuous variation. Most questions still dealt with static unknown quantities or, if they involved variation at all, involved the “chunky” conceptualization of variation described by Castillo-Garsow et al. (2013). While industrial needs had pushed students in these decades towards understanding fractions as abstract numbers on a continuous number line, the mathematical thinking exhibited on the exams still remained largely constrained by the technological affordances and demands of the pre-computer age.

CONCLUSION AND IMPLICATIONS FOR FUTURE RESEARCH

Preliminary analysis of New York state exams suggests that broad societal transformations can spur changes in how people engage with mathematical concepts. In early twentieth-century New York, evidence indicates that the shift from rural to urban settings required engagement with new technology, organizational systems, and economic relationships, which may have pushed people toward more sophisticated ways of thinking about fractions. One direction for future research would be to focus on today’s cultural and technological contexts and the demands they create for mathematical understanding. Siegler & Ramani’s (2008) study on the connection between board games and mathematical understanding is one example, but the potential is much broader. The ubiquitous battery percentage bar on the face of a cell phone, for example, may affect the way middle school children understand percentages. Further research examining such cultural tools could have important implications for how we teach these mathematical concepts.

References

- Ahl, L. M., & Helenius, O. (2024). Is it a fraction, or shall I divide it? In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.), *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 17-24). PME.
- Castillo-Garsow, C., Johnson, H. L., & Moore, K. (2013). Chunky and smooth images of change. *For the Learning of Mathematics*, 33(3), 31-37.
- González-Forte, J. M., Fernández, C., Van Hoof, J., & Van Dooren, W. (2021). From discreteness to infinity: Stages in students' understanding of the rational number density. In M. Inprasitha, N. Changsri, & N. Boonsena (Eds.), *Proceedings of the 44th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 337-344). PME.
- Jefferson Math Project. (1866-2024). New York State State Regents Examinations archive [Digital archive]. www.jmap.org

- Mack, N. K. (1993). Learning rational numbers with understanding: The case of informal knowledge. In Carpenter, T.P., Fennema, E., & Romberg, T.A. (Eds.). (1993). *Rational Numbers: An Integration of Research* (1st ed.). Routledge.
- McMullen, J., Laakkonen, E., Hannula-Sormunen, M., & Lehtinen, E. (2015). Modeling the developmental trajectories of rational number concept(s). *Learning and Instruction*, 37, 14-20.
- Ni, Y., & Zhou, Y.-D. (2005). Teaching and learning fraction and rational numbers: The origins and implications of whole number bias. *Educational Psychologist*, 40(1), 27-52.
- Schneider, M., & Siegler, R. S. (2010). Representations of the magnitudes of fractions. *Journal of Experimental Psychology: Human Perception and Performance*, 36(5), 1227-1238.
- Siegler, R. S., & Ramani, G. B. (2008). Playing linear number board games promotes low-income children's numerical development. *Developmental Science*, 11(5), 655-661.
- Siegler, R. S., Fazio, L. K., Bailey, D. H., & Zhou, X. (2013). Fractions: The new frontier for theories of numerical development. *Trends in Cognitive Sciences*, 17(1), 13-19.
- Siegler, R. S., Thompson, C. A., & Schneider, M. (2011). An integrated theory of whole number and fractions development. *Cognitive Psychology*, 62(4), 273-296.
- Sirotic, N., & Zazkis, R. (2007). Irrational Numbers: The Gap between Formal and Intuitive Knowledge. *Educational Studies in Mathematics*, 65, 49-76.
- Smith, D. E. (1909). *The teaching of arithmetic*. Ginn and Company.
- State Education Department. (1965). *100 years of quality control in education: Regents examinations 1865-1965*. The University of the State of XXXX, Regents Examination and Scholarship Center.
- Stigler, J. W., Givvin, K. B., & Thompson, B. J. (2010). What community college developmental mathematics students understand about mathematics. *The MathAMATYC Educator*, 10(4), 4-16.
- XXXX Census Bureau. (1901). Part 1: Introduction, population of states and territories, density of population, center of population and its median point, population of counties and incorporated places, urban and rural population, sex, general nativity. In *Twelfth census of the XXXX, taken in the year 1900*.
- Vamvakoussi, X., & Vosniadou, S. (2004). Understanding the structure of the set of rational numbers: A conceptual change approach. *Learning and Instruction*, 14(5), 453-467.
- Vamvakoussi, X., & Vosniadou, S. (2010). How many decimals are there between two fractions? Aspects of secondary school students' understanding of rational numbers and their notation. *Cognition and Instruction*, 28(2), 181-209.



MAKING SURE THAT MATHEMATICS EDUCATION
RESEARCH REACHES THE CLASSROOM



Universidad
de Chile

LOH
Universidad
de O'Higgins