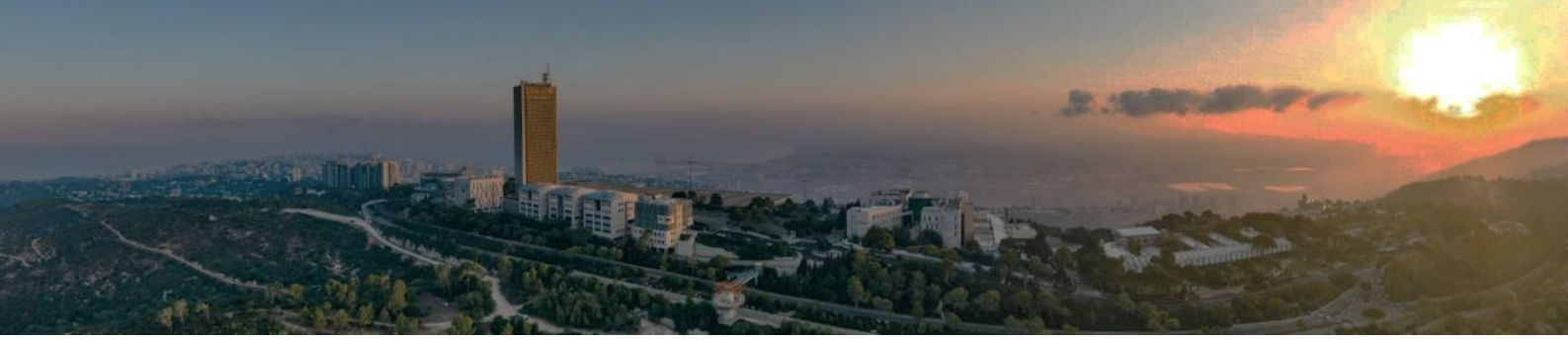


PROCEEDINGS

OF THE

46th CONFERENCE

of the International Group for the Psychology of Mathematics Education

Haifa, Israel | July 16 – 21, 2023

PME-46

MATHEMATICS EDUCATION FOR GLOBAL SUSTAINABILITY

Volume 4

Editors: Michal Ayalon, Boris Koichu, Roza Leikin, Laurie Rubel and Michal Tabach

PROCEEDINGS

of the

46th CONFERENCE

of the International Group for the Psychology of Mathematics Education

Haifa, Israel

July 16 – 21, 2023

Editors:

Michal Ayalon, Boris Koichu, Roza Leikin, Laurie Rubel and Michal Tabach

Volume 4

Research Reports

P – Z

PME-46
MATHEMATICS EDUCATION
FOR GLOBAL SUSTAINABILITY

Cite as:
Michal Ayalon, Boris Koichu, Roza Leikin, Laurie Rubel & Michal Tabach (Eds.). (2023) Proceeding of the 46th Conference of the International Group for the Psychology of Mathematics Education (Vol. 4). University of Haifa, Israel: PME

Website: https://pme46.edu.haifa.ac.il/
The proceedings are also available via
https://www.igpme.org/publications/current-proceedings/

ISSN: 0771-100X
ISBN: 978-965-93112-4-8

Printed by the University of Haifa

Logo Design: Sharon Spitz (https://www.sharonspitz.com/)
Cover design: Sharona Gil (Facuty of Education, University of Haifa)

TABLE OF CONTENTS VOLUME 4

RESEARCH REPORTS (P-Z)

TOWARDS A SPECIFICATION OF DIGITAL COMPETENCES FOR STEM TEACHERS IN AN EDUCATIONAL CONTEXT. ELICITING EXPERTS' VIEWS 4-3

Rouven Pankrath and Anke Lindmeier

STUDENTS' USE OF UNNECESSARY BRACKETS AS A WAY OF EXHIBITING STRUCTURE-SENSE 4-11

Ioannis Papadopoulos and Athina Thoma

TYPES AND FEATURES OF DIALOGICAL TASKS FROM MATHEMATICS TEACHERS' PERSPECTIVE 4-19

Reut Parasha and Boris Koichu

ARE EXPERTS' NOTICING FOCUSES REGARDING THE LEARNING POTENTIAL OF TASKS AND ITS USE CONSISTENT ACROSS INSTRUCTIONAL SITUATIONS? A SECONDARY ANALYSIS 4-27

Josephine F. Paul, Anika Dreher, Ting-Ying Wang, Feng-Jui Hsieh, Linn Hansen, Anke Lindmeier

GENDER-RELATED BELIEFS OF PROSPECTIVE MATHEMATICS TEACHERS 4-35

Georg Pfeiffer and Daniela Assmus

REPRESENTING COVARIATION FUNCTIONAL SITUATIONS IN A TABLET-ENABLED DIGITAL LEARNING ENVIROMENT 4-43

Marios Pittalis, Eleni Demosthenous and Ute Sproesser

A META-DISCIPLINARY REFLECTION ON A STEAM SCHOOL ACTIVITY: THE ROLE OF MATHEMATICS 4-51

Gabriella Pocalana, Ornella Robutti and Giulia Bini

UNIT STRUCTURES RARELY ARTICULATED: TEACHERS' EXPLANATIONS OF MEANINGS OF MULTIPLICATION 4-59

Susanne Prediger[1,2] and Anke Wischgoll

STUDYING THE ROLE OF PSEUDO-OBJECTS IN PROOF BY CONTRADICTION 4-67

Kostas Probonas and Giorgos Psycharis

LEVELS OF MATHEMATICAL KNOWLEDGE IN LINEAR ALGEBRA FOR ENTERING UNIVERSITY 4-75

Kolja Pustelnik, Stefanie Rach, Daniel Sommerhoff[2], and Stefan Ufer

PROSPECTIVE PRIMARY TEACHERS' UNDERSTANDING OF ONE-DIMENSIONAL PHENOMENA: LINE, RAY AND SECTION 4-83

Simone Reinhold and Bernd Wollring

THE ROLE OF LANGUAGE-AS-RESOURCE AND LANGUAGE-AS-POLITICAL IN COLLEGE MATHEMATICS COURSES 4-91

Jocelyn Rios

EMBODIED CURIOSITY: A FRAMEWORK FOR MATHEMATICAL MEANING-MAKING 4-99

Sheree Rodney

IHIGH SCHOOL STUDENTS' PERCEPTIONS OF THE RELEVANCE OF MATHEMATICS IN HIGHER EDUCATION 4-107

Dunja Rohenroth, Irene Neumann and Aiso Heinze

INTRODUCTION AND THEORETICAL BACKROUND 4-107

Joshua M. Ruk and Laura R. Van Zoest

TEACHER CHANGE AND INCLUSIVE INTERVENTIONS FOR LEARNERS WITH MATHEMATICS DIFFICULTIES 4-123

Robyn Ruttenberg-Rozen and Marc Husband

WHAT IS A "GOOD" ARGUMENTATION IN MATHEMATICS CLASSROOM? 4-131

Saccoletto Marta

DIDACTIC-MATHEMATIC KNOWLEDGE TRAITS OF PRE-SERVICE TEACHERS WHEN POSING AND SOLVING ROBOTIC PROBLEMS 4-139

Gemma Sala-Sebastià, Adriana Breda, Alicia Sánchez and Vicenç Font

INDIVIDUAL CONCEPTION FRAMES AS A CONCEPT FOR THE ANALYSIS OF MATHEMATICAL LEARNING 4-147

Alexander Salle and Marcus Schütte

DIGITAL MONITORING OF FRACTION LEARNING: ADAPTING A TEST FOR KNOWLEDGE OF FRACTION SUBCONSTRUCTS 4-155

Constanze Schadl, Anke Lindmeier

CHARACTERIZING EXTERNAL VISUALIZATION INTERVENTIONS: A SYSTEMATIC LITERATURE REVIEW 4-163

Johanna Schoenherr and Stanislaw Schukajlow

TEACHERS' DIAGNOSTIC ACTIVITIES DURING TASK-BASED ASSESSMENTS IN A DIGITAL SIMULATION 4-171

Christian Schons, Andreas Obersteiner, Frank Fischer and Kristina M. Reiss

STATISTICAL THINKING AND VIEWING PATTERNS WHEN COMPARING DATA DISTRIBUTIONS: AN EYE-TRACKING STUDY WITH 6TH AND 8TH GRADERS 4-179

Saskia Schreiter and Markus Vogel

HOW DO MATHEMATICS TEACHERS LEARN TO CREATE A MATHEMATICAL STORYLINE IN PROBLEM-BASED LESSONS? 4-187

Gil Schwarts, Patricio Herbst and Amanda Brown

MATHEMATICAL REASONING TYPES AS GENDERED? VIEWS FROM PALESTINIAN/ARAB ISRAELI TEACHERS 4-195

Juhaina Awawdeh Shahbari, Laurie Rubel and Fatema Kabha

DYNAMIC VISUALIZATION AND EMBODIED DESIGN FOR TRIGONOMETRY LEARNING: LOOKING OR DOING? 4-203

Anna Shvarts and Linda Zenger

STRATEGY USE IN NUMBER LINE TASKS OF STUDENTS WITH AND WITHOUT MATHEMATICAL DIFFICULTIES: A STUDY USING EYE TRACKING AND AI 4-211

Anna Lisa Simon, Parviz Asghari, Achim J. Lilienthal, & Maike Schindler

A NOVICE TEACHER'S IDENTITIES – FROM LOSING HER BALANCE TO REGAINING HER CONFIDENCE 4-219

Charlotte Krog Skott and Jeppe Skott

DIDACTIC SUITABILITY CRITERIA IN TEACHERS' PRACTICAL ARGUMENTATION IN THE PHASE OF DESIGN OF A LESSON STUDY CYCLE ABOUT FUNCTIONS 4-227

Telesforo Sol[1], Alicia Sánchez, Adriana Breda and Vicenç Font

FOSTERING STUDENTS' KNOWLEDGE ABOUT PROOF 4-235

Femke Sporn, Daniel Sommerhoff and Aiso Heinze

FROM 2D TO 3D: SUPPORT of a 3-Dimensional DYNAMIC GEOMETRY ENVIRONMENT IN LEARNING proof 4-243

Camilo Sua, Angel Gutiérrez and Adela Jaime

CHINESE LANGUAGE LEARNERS' READING COMPREHENSION WHEN SOLVING MATHEMATICAL WORD PROBLEMS 4-251

Emily S. W. Sum, Miranda, K. Y. Wong, Antonia, Y. T. Yip and Wee Tiong Seah

ENHANCING SPATIAL REASONING THROUGH GEOMETRY TRANSFORMATION INSTRUCTION IN GHANA 4-259

Mawuli Kofi Tay and Armando Paulino Preciado Babb

CONNECTING MATHEMATICS LEARNING TO LEARNING ABOUT STRUCTURAL RACISM IN THE UNITED STATES 4-267

Eva Thanheiser and Molly Robinson

APPLETS AND PAPER & PENCIL TASKS AS RESOURCES FOR WORKING WITH MATHEMATICAL REPRESENTATIONS 4-275

Odelya Tzayada and Michal Tabach

GROUNDING CHINESE NEW STANDARDS' FOCUS ON COUNTING-UNITS IN A CONSTRUCTIVIST, UNITS-AND-OPERATIONS MODEL 4-283

Ron Tzur, Rui Ding, Yunpeng Ma, Rongjin Huang and Bingqian Wei

ANALYSING THE QUALITY OF ADVANCED MATHEMATICS LECTURES REGARDING THE PRESENTATION OF THEOREMS AND PROOFS – THE CASE OF REAL ANALYSIS LECTURES 4-291

Karyna Umgelter and Sebastian Geisler

STRATEGY USE IN NUMBER LINE ESTIMATIONS OF FRACTIONS – AN EXPLORATORY STUDY IN SEARCH FOR ADAPTIVE EXPERTISE 4-299

Wim Van Dooren

MAKING SENSE OF ZERO TO MAKE SENSE OF NEGATIVE NUMBERS 4-307

Joëlle Vlassis and Isabelle Demonty

THE DISCOURSE MAPPING TREE AS A TOOL FOR ANALYZING THE POTENTIAL AND IMPLEMENTATION OF LINEAR ALGEBRA TASKS 4-315

Miriam N. Wallach, Einat Heyd-Metzuyanim and Ram Band

ANSWER PATTERNS OF JAPANESE SECONDARY SCHOOL STUDENTS IN TIMSS 2015 MATHEMATICS SURVEY 4-323

Koji Watanabe

REPRESENTING 'TALL AND SHORT' IN DRAWINGS – PRE-SCHOOL TO YEAR 2 4-331

Jennifer Way

IAPPLYING A COMMOGNITIVE-BASED FRAMEWORK TO PROMOTE TEACHERS' COMMUNICATION ABOUT REASONING AND PROVING 4-339

Merav Weingarden and Orly Buchbinder

FROM UNIVERSITY TO SCHOOL: EXPLORING BEGINNING TEACHERS INTEGRATING REASONING AND PROVING 4-347

Merav Weingarden and Orly Buchbinder

LEARNING ABOUT DIGITAL TECHNOLOGIES OF THE WORKING WORLD IN REGULAR MATH CLASSES? TEACHING COMPOSITE BODIES WITH 3D PRINT AS A LEARNING CONTEXT 4-355

Mira H. Wulff, Anika Radkowitsch and Aiso Heinze

TEACHERS' MULTIPLE AND ADAPTIVE NOTICING DRIVEN BY THEIR FRAMING OF PROFESSIONAL OBLIGATIONS IN THE CONTEXT OF A PROVING ACTIVITY 4-363

Mei Yang, Andreas J. Stylianides and Mateja Jamnik

ELEMENTARY PRESERVICE TEACHERS' NOTICING OF EXEMPLARY LESSONS: A COMPARISON OF NOTICING FRAMEWORKS 4-371

Qiaoping Zhang, Yicheng Wei and Jing Liang

EXPLORING THE ROLE OF PEDAGOGY IN MATHEMATICAL CREATIVITY VIA MULTIPLE SOLUTION TASKS: A COMPARATIVE STUDY OF TWO SCHOOLS IN CHINA 4-379

Ying Zhang

SIXTH GRADERS' LEARNING OF MULTIPLICATIVE STRUCTURE PROBLEMS THROUGH THE VARIATION PRINCIPLE 4-387

Cristina Zorrilla, Ceneida Fernández, Anna-Katharina Roos, Salvador Llinares, and Susanne Prediger

VOLUME 4

RESEARCH REPORTS

P - Z

PME-46

MATHEMATICS EDUCATION
FOR GLOBAL SUSTAINABILITY

TOWARDS A SPECIFICATION OF DIGITAL COMPETENCES FOR STEM TEACHERS IN AN EDUCATIONAL CONTEXT. ELICITING EXPERTS' VIEWS

Rouven Pankrath and Anke Lindmeier

FSU Jena, Germany

Despite an increasing number of available frameworks for (future) teachers' digital competences, it often remains unspecified what teachers should know and be able to do. Hence, deciding on the focus of courses is still challenging. We initiated a Delphi process with stakeholders from research, school administration and practice in a local educational context to identify digital competences central for STEM teachers. This report covers the first stage of the process, where competence expectations synthesized from different frameworks were subjected to relevance evaluations. The results indicate a high degree of consensus among the experts, and experts from various fields of expertise differ only in a few aspects. We discuss how the process may inform others challenged to decide on questions related to (future) teacher education.

INTRODUCTION

Across the world, national policy actions stress the importance of preparing (future) teachers for working in a digital world. Hence, digital competences are seen as an essential aspect of the professional competence of teachers that enable teachers to use digital technologies in and for teaching. For instance, they should be able to integrate digital technologies effectively into teaching processes and use them for lesson preparation or communication with parents.

Several international (and also national) frameworks conceptualize digital competences or describe their range, for instance, the TPACK model (Koehler & Mishra, 2009) and the DigCompEdu (Redecker, 2017). However, existing frameworks for digital competences for teachers are hardly suited to decide what should be first and foremost targeted in courses fostering digital competence for (future) teachers of specific subjects. So, in many educational contexts, educators need to know what might be considered relevant by others holding responsibilities in the same context. This contribution addresses the problem for STEM teachers for (upper) secondary level in a first step by investigating whether it is possible to elicit a consensus of different stakeholders from research, school administration, and practice within an educational context regarding what might be relevant digital competences for all teachers.

THEORETICAL BACKGROUND

Digital competences of teachers can be described as a set of knowledge, skills, and attitudes related to the use of digital technology in education (Ferrari, 2012), and, accordingly, frameworks describe them in different ways. For example, the TPACK framework by Koehler and Mishra (2009) portrays teachers' professional knowledge

2023. 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. 3-10). PME 46.

as an overlap of technological, pedagogical, and content knowledge. This model is compact yet rather abstract and focuses on critical areas of teachers' knowledge, four related to digital competences. Another widespread framework for teachers' digital competences is DigCompEdu (Redecker, 2017). It comprehensively lists 22 different abstract competences in six areas. For example, teachers are expected to be able to use digital resources for teaching by creating and modifying them (2.1) or for professional collaboration (1.2), but they are also expected to support the students' responsible use of technology (6.4). Every competence is further characterized by hierarchical proficiency expectations in eight levels. Although this comprehensive framework is less abstract than TPACK, it still describes competences in a general manner and comprises a wide variety of different expectations.

Typically, digital competence expectations are also documented by national policies. In the federal state of Germany, for instance, the standing conference of ministries of education (KMK, 2017) issued competence expectations built on the cited frameworks and other resources. For instance, they incorporate ideas of critical education, which are often referenced in general education and its sciences, but not in (European) mathematics education (Skovsmose, 1985). However, the national policy documents also have commonalities with the international frameworks, like being very abstract.

As overarching documents, all frameworks are also limited informative regarding the expectations for teachers of specific subjects, for example, mathematics. At the same time, study results indicate that the use of digital technology differs between subjects, for example, mathematics and natural sciences (Mullis et al., 2020), so expectations may also have to be differentiated according to the subjects taught.

Moreover, the frameworks often lack information regarding an important question for the purpose of teacher education and training: What might be considered a minimal set of competences relevant for all (future) teachers? It can be assumed that even experts would answer the question differently, as teacher education and training are fields of shared responsibilities and partially disconnected (scientific) discourses.

RESEARCH QUESTIONS

This contribution addresses the problem of a need for more specific learning expectations regarding digital competences for (future) teachers. As described, current models of teacher digital competences do not guide what might be particularly relevant for teachers. They are also less informative regarding specific subjects, despite there is evidence that the use of digital technology differs between subjects. Moreover, general education reportedly uses different frameworks than subject-specific educational research, but there is a lack of evidence on whether the views of experts from different fields differ. According to our interest, we focus on STEM subjects which are often referenced as being at the fore of using digital technology.

We aim to answer the following research questions:

RQ1) Is there a consensus regarding which digital competences are relevant for all mathematics and science teachers by experts within a certain local context?

RQ2) Do experts with different fields of expertise (e.g., in the subjects, in general education) have different views on the relevance of competences?

DESIGN OF THE STUDY AND METHODS

The study is part of a larger study implementing the Delphi method (Diamond et al., 2014), which can be described as a moderated collaborative problem-solving process. This report covers the preliminary results of the first round, where we investigated whether experts' views on digital competence statements reached a consensus. As there were already frameworks for digital competences referenced in our educational context, we decided to start with a set of statements based on an analysis and synthesis of the frameworks (see below) in a structured online questionnaire. In this contribution, we report on the procedure and the results of this online questionnaire.

We expected the online questionnaire to be suited to identify statements seen consensually as relevant and others as not. The results should be used in the future second Delphi round to further specify learning expectations according to the identified relevant competence statements in group discussions with the experts. This should finally allow deciding on the design of courses for (future) STEM teachers fostering digital competences considered relevant by stakeholders with shared responsibilities in teacher education and training in our educational context.

Design of the Instrument

To identify a set of statements to be used in our questionnaire, we started analyzing the structure of an online course called *digi4all* (Seegerer et al., 2021), designed to equip pre-service and in-service teachers with basic digital concepts. As Seegerer created this course in collaboration with other subject education scientists through a similar consensus process, we expected the course to be a good starting point, yet being coined from the perspective of the author, a computer science education expert. We further subjected the frameworks of TPACK, DigCompEdu, and the relevant national educational policies, including the local curriculum for the school subject of computer science (indicating the local expectations of general education outcomes related to digital competences) to a qualitative content analysis. We synthesized the expectations of the different sources into statements and grouped the statements according to categories. Therefore, we merged and expanded the categories of *digi4all*, and the process resulted in eight competence areas (Table 1). When writing the statements, we aimed at easy readability of the statements and provided examples in case of possible divergent understanding of terms.

We structured the statements in each competence area in two parts differentiating between knowledge and skills. So, the statements in the first part referred to topics that may be expected to be known, whereas the statements in the second part referred to

actions that may be expected to be mastered. To illustrate this, we present one statement referring to knowledge (indicated by K in the label) and one to skills (S) for the competence areas CA1, CA3, and CA6: “Binary number system as the basis of ‘digital functioning’” (CA1Kbin); “Operating with binary digits, e.g., conversion of binary to decimal numbers, binary addition, subtraction” (CA1Sbin); “Fundamentals of statistics” (CA3Kstat); “Evaluate collected information and data and present it appropriately for the addressee” (CA3Sstat); CA6: “Psychological effects of social networks (e.g., cyberbullying)” (CA6Ksocnet); “Using video conferencing tools” (CA6Svidcon).

Label	Content [number of statements]	Label	Content [number of statements]
CA1	Fundamentals of the functionality and use of a computer [14]	CA5	Fundamentals of media culture and influence of media on daily life [6]
CA2	Fundamentals of the functionality and use of the internet [10]	CA6	Communicating through and collaborating with digital technologies [10]
CA3	Getting, saving and evaluating data and information [11]	CA7	Designing digital learning environments (in general) [8]
CA4	Understanding, using and evaluating algorithms [8]	CA8	Using and evaluating subject-specific digital tools [8]

Table 1: Overview of the eight competence areas (abbr. CA) with a short description of contents and the number of statements identified.

To elicit what the experts considered to be digital competences relevant for all mathematics and science teachers, we asked them to rate each of the statements. The experts could rate the statement as being relevant for all teachers or not. They could further indicate whether they considered a basic level or an advanced level of knowledge/skills as necessary, which was asked to inform later stages of the Delphi process. In addition, they could decide not to rate the statement if they feel to do so (“This is not something I can assess”).

For each CA, the experts could comment in free text fields if they had suggestions for competences that are needed by all teachers but not covered by the presented statements. It was possible for the experts to navigate freely through the questionnaire at any time and to change the given answers.

Sampling method and sample

The experts for this questionnaire were intentionally sampled. Our educational context refers to a German federal state (Thuringia) and the Gymnasium level. All future secondary teachers for this level are educated at one university. Since our focus is on mathematics and science, we contacted all STEM education professors. In addition, we also contacted professors in the field of educational sciences that are responsible for the general education parts of teacher education. Finally, we contacted the ministry of education of the federal state as representative of the educational administration. In each case, we contacted the head of the department. We asked for personal participation and the nomination of a given number of post-/doctoral researchers or

employees with relevant expertise to participate in the questionnaire. In addition, we asked to nominate up to two expert teachers experienced in mentoring future teachers and one subject expert (professor) experienced in teaching future teachers for each subject. Overall, we aimed at a list of 5 to 7 participants per field of expertise. With this sampling strategy, we intended to mirror the shared responsibilities in teacher education and training in our educational context.

To investigate RQ2 regarding potential differences according to the experts' field of expertise, we build groups as follows: Mathematics and computer science (group 1), biology, physics, and chemistry (group 2), educational sciences and educational administration (group 3). Whereas group 2 refers to the natural sciences, group 1 spans computer science and mathematics as subjects with a common root and (still) partly common grounds in mathematics. Group 3 represents experts that have, by nature of their professional field, a general perspective on teachers and their competences independent of the subjects taught.

In the end, we invited 46 people to take part in our questionnaire as experts. We informed the experts about the goals and procedures and whether they were free to participate. The experts were informed that we could identify responses with personal information, which is necessary for the next round of the Delphi process. This research report is a working report by the time of January 2023. At this time, we received complete responses from 36 of the invited experts (age: M = 43.7; SD = 12.4).

Group	Field of expertise	Number of responses	Response rate
1 (n=14)	Mathematics (MA)	6	86%
	Computer Science (CS)	8	89%
2 (n=11)	Biology (BI)	4	80%
	Physics (PH)	4	80%
	Chemistry (CH)	3	60%
3 (n=11)	Educational Sciences (EdS)	6	60%
	Educational Administration (EdA)	5	100%
	Total	36	78%

Table 2: Number of participants by field of expertise and response rates.

Data analysis

We report the experts' rating of the full set of 75 competence statements structured in eight competences areas. We treated answers where experts decided not to rate the statement as missing. To answer RQ1, we counted whether experts considered each statement as a relevant requirement for all teachers without differentiating between possibly different levels of sophistication the experts might expect (basic vs. advanced) statements seen by at least 75% of the experts as being relevant were considered consensually representing a relevant expectation (typical Delphi criterion, Diamond et al., 2015). Some experts provided additional comments in the text fields provided with each competence area (98 written comments in total). The detailed analysis of these comments cannot be part of this working report, but a first inspection led to the

impression that the responses were mainly comments on the presented statements with few suggestions for additions.

Regarding RQ2, we subjected differences in agreement rates between the groups of experts with different fields of expertise to a *Chi*²-test of independence and manually inspected observed differences.

RESULTS

General findings will first be reported aggregated for the competence areas. With our criterion of 75% necessary agreement rate, we note a trend over the different competence areas. As CA1 to CA4 are more about the technical and general aspects of using technology and five to eight are more about teaching-specific aspects, nearly all statements in the competence areas five to eight exceeded the 75% agreement criterion. For example, the statements CA6Ksocnet and CA6Svidcon (see above) reached 100% agreement rates. In contrast, several statements did not reach the specified agreement rate in CA1 to CA4. For example, the statement CA1Kbin reached 52%, and CA1Sbin only 13% agreement rate. However, overall relevance agreement rates were high (M = 80%, SD = 23% agreement rate). To answer RQ1, the experts' ratings indicate that 59 out of the 75 presented statements were consensually considered to be relevant for all (future) STEM teachers, at least on a basic level.

CA 1	CA 2	CA 3	CA 4	CA 5	CA 6	CA 7	CA8
10 of 14	5 of 10	9 of 11	3 of 8	6 of 6	10 of 10	8 of 8	8 of 8

Table 3: Number of statements with an agreement rate at or above 75% across all experts (N=36) in each competence area.

Regarding RQ2, we exemplarily focus in this report on CA1 and CA3, as the experts' responses show certain variations in these areas. The results of the *Chi*²-test indicate that only for 2 of the 25 statements, there are significant differences between the agreement rates. To illustrate how the experts' views differ in our study, we present details for selected statements. Among the statements in CA1, we presented two statements referring to skills in using software for text editing and presentation. The experts consensually rated these as relevant for all teachers (100% agreement rate).

	Overall	Group 1 MA + CS	Group 2 BI + CH +PH	Group 3 EdS + EdA	*Chi*²-test
CA1Kbin	52%	29%	44%	90%	*
CA1Sbin	13%	7%	11%	22%	n.s.
CA3Kstat	61%	67%	67%	50%	n.s.
CA3Sstat	100%	100%	100%	100%	n.s.

Table 4: Agreement rates split up by groups on selected statements of CA1 and CA3 to illustrate how experts' responses vary (* < .05, n.s. not significant).

In contrast, the two CA1 statements about binary numbers given above achieved varying agreement rates (Table 4), with the agreement of group 1 lower than group 2 and the agreement of group 2 lower than group 3. But, the descriptive differences correspond only for CA1Kbin to a significant difference. As other examples, in CA3,

we presented two statements about statistics, one referring to knowledge (CA3Kstat) and one to skills (CA3Sstat). Despite certain variances in agreement rates between groups (Table 4), the differences cannot be considered statistically significant.

To sum up, and as already indicated by the high consensus rates reported above, the views of experts with different fields of expertise only differ in a few cases. It has to be noted that we exemplarily focused on CA1 and CA3 in this preliminary report, and the analyses of the other two areas showing certain variations are pending.

DISCUSSION

This contribution reports the results of a study investigating whether experts of different fields of expertise in an educational context have similar views on what is considered digital competences relevant for all STEM teachers. As a starting point, we presented a set of 75 statements in eight competence areas synthesized from different frameworks. The results indicate that the experts consensually rate a wide range of statements as relevant for all teachers. We see indications that the consensus is almost perfect when expectations are particularly teaching-specific, as in competence areas “fundamentals of media culture and influence of media on daily life” (CA5), “communicating through and collaborating with digital technologies” (CA6), “designing digital learning environments (in general)” (CA7) and “using and evaluating subject-specific digital tools” (CA8). For less teaching-specific areas, experts’ ratings are more differentiated so that, for example, general skills like using standard software for text editing are undisputedly seen as relevant. At the same time, questions related to the hidden principles of technologies, like binary numbers or statistical principles, did not reach a consensus in our study.

Against expectations, the results of our preliminary analysis suggest that differences between the groups of experts from different fields of expertise are not very salient. We illustrated this by the examples of the statements referring to the binary system and its operations. From a mathematical point of view, it is remarkable that the relevance agreement rates are generally low but lowest for the group of experts from mathematics and computer science and highest for the educational sciences and administration group, which seems paradoxical at first sight. One possible explanation might be that experts more familiar with these concepts underestimate their relevance. We will use the planned group discussion to elicit the reasonings behind the relevance ratings by experts from different fields for possible explanations.

Our study certainly has limitations. First, we must remember that we focus on a certain educational context. Hence, the findings are not generalizable across contexts. However, to our knowledge, studies that systematically investigate whether experts with shared responsibility in teacher education and training have similar or different views on digital competence expectations are rare. Our methods might also inform other studies addressing similar teacher education problems. Second, so far, our competence statements are still abstract and may be interpreted differently by experts with different backgrounds. In addition, we did not consider whether the experts' views

might show considerable variability regarding the expected level of the competences, even if a statement was consensually rated as relevant. We will focus on these aspects in further analyses of the data and the second round of the Delphi process, where we intend to initiate group discussions and aim for a consensus on the level of specific expectations of mastery. The process might still show that, despite a perfect agreement regarding the competence statements (e.g., teachers have to know about the "Functioning of social networks", CA6Kfunsoc), experts with different backgrounds, like mathematics educators and general educators, may mean something different by this statement.

On the one hand, our findings underline the relevance of a wide variety of digital competences for (future) STEM teachers. On the other hand, our study shows that the original problem of deciding what should be covered by teacher education courses remains even though we applied a strong consensus criterion. This is remarkable, given teacher education and training experts' diverse backgrounds. However, it also suggests that the different stakeholders with shared responsibilities in teacher education and training might succeed in initiating a common discourse.

ACKNOWLEDGEMENTS

This study is part of THInKI, founded by the BMBF (Federal Ministry of Education and Research) with the registration number 16DHBKI084.

REFERENCES

Diamond, I. R., Grant, R. C., Feldman, B. M., Pencharz, P. B., Ling, S. C., Moore, A. M. & Wales, P. W. (2014). Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. Journal of clinical epidemiology, 67(4), 401–409.

Ferrari, A. (2012). Digital Competence in Practice: An Analysis of Frameworks. JRC68116.

KMK. (2017). Bildung in der digitalen Welt. [Standing conference of the ministries of education, Germany].

Koehler, M. J. & Mishra, P. (2009). What Is Technological Pedagogical Content Knowledge (TPACK)? Contemporary Issues in Technology and Teacher Education, 9(1), 60–70.

Mullis, I. V. S., Martin, M. O., Foy, P., Kelly, D. L. & Fishbein, B. (2020). TIMSS 2019: International Results in Mathematics and Science.

Redecker, C. (2017). European Framework for the Digital Competence of Educators: DigCompEdu. JRC107466.

Seegerer, S., Michaeli, T. & Romeike, R. (2021). Informatische Grundlagen in der allgemeinen Lehrkräftebildung: Erkenntnisse und Erfahrungen aus einem online-gestützten Studienangebot. In L. Humbert (Hrsg.), Informatik-Bildung von Lehrkräften in allen Phasen. Lecture Notes in Informatics, pp. 153–162.

Skovsmose, O. (1985). Mathematical Education versus Critical Education. Educational Studies in Mathematics, 16(4), 337–354.

Vuorikari, R., Kluzer, S. & Punie, Y. (2022). DigComp 2.2: The Digital Competence Framework for Citizens. JRC128415.

STUDENTS' USE OF UNNECESSARY BRACKETS AS A WAY OF EXHIBITING STRUCTURE-SENSE

Ioannis Papadopoulos[1] and Athina Thoma[2]

Aristotle University of Thessaloniki[1], University of Southampton[2]

In this paper, we analyze high school students' answers to algebraic questions focusing on their use of unnecessary brackets. This is a part of a larger study exploring students' use and non-use of brackets at various educational levels and on different mathematical topics. Our focus in this paper is on the instances of the use of unnecessary brackets and their connection to students' structure sense. The analysis of the students' written solutions showed three main categories regarding keeping the substitution of numerical or algebraic expressions separate, illustrating the application of a rule focusing on explicitly illustrating the different elements, and showing the grouping of terms in order to assist with the following evaluations.

INTRODUCTION

Brackets constitute a structural element of arithmetic and algebraic expressions since they determine the relations that exist between different parts of these expressions. For example, in the equality $\frac{a+b}{c+d} = (a+b) \div (c+d)$ brackets preserve the structure of the initial expression by making clear the relation between the numerator and the denominator of the fraction. Despite their importance in the structure of arithmetic and algebraic expressions, it seems that sometimes students do not physically write the necessary brackets. So, the right part of the equation above might be written as $a + b \div c + d$. In this case, students very often evaluate the expressions in a way that mirrors the presence of brackets, acting thus as if the brackets were there. Papadopoulos and Gunnarsson (2020) refer to this phenomenon as the 'use of mental brackets' and claim that their use is connected to students' structural understanding. However, there is another option beyond the dichotomy 'use or non-use of necessary brackets'. This is the use of unnecessary brackets. Our research focuses on whether, and if yes to what extent, the use of unnecessary brackets relates to students' structure sense.

THEORETICAL BACKGROUND

For most students, brackets are related to rules for the order of operations. This is reasonable since in school brackets are introduced and used mainly in the following context. When the student is asked to evaluate an expression containing bracketed sub-expressions, then the content of these brackets takes precedence. Thus, brackets determine which operation should be calculated first. The emphasis on this aspect of brackets underplays the role of brackets as structural elements of the expression and therefore the significance of students' structure sense. Structure sense is a term introduced initially by Linchevski and Livneh (1999) aiming to describe students'

2023. 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. 11-18). PME 46.

difficulties when they use their knowledge of arithmetic structure during their first steps of learning algebra. According to them, structure sense for arithmetic expressions is related to students' ability to correctly parse the expression and identify the relations between the components of the expression as well as between the components and the whole. The term was refined later by Hoch and Dreyfus (2006) who claim that students exhibit structure sense when they are able to recognize a familiar structure in its simplest form; consider a compound term as a single entity; and, by substituting appropriately they identify a familiar structure in a more complex form and make use of the expression's structure through appropriate manipulation. Hoch and Dreyfus (2004) found that the presence of brackets plays a significant role in the way students use the structure of the expression. More specifically, they found that the presence of brackets helps students focus their attention on certain terms of the expression thus influencing positively their structure sense.

On the antipode, Papadopoulos and Gunnarsson (2020) working with primary school students examined the absence of brackets and its connection with structure sense. They investigated Grade 5 and 6 students on the evaluation of rational arithmetical expressions. The students had to initially write these expressions in their horizontal form and then evaluate them. In this case, the use of brackets is necessary to preserve the structure of the given rational expression. The analysis illustrated that many students did not physically write brackets in the horizontal form but evaluated the expressions as if the brackets were there. From the technical point of view, the way they evaluated the horizontal expression violated the rules for the order of operations but in essence, they respected the structure of the initial expression. Papadopoulos and Gunnarsson (2020) refer to this phenomenon using the term 'mental brackets'. So, the use of mental brackets seems to have assisted students in perceiving the expressions in the way they should be perceived regardless of students' writing. Later, Papadopoulos and Thoma (2022) investigated the same issue in the solutions of high school students in the context of algebra. They found that students use mental brackets mainly when substituting values in variables. This use of mental brackets was identified as either applied to whole expressions (global mental brackets) or just to some of their parts (local mental brackets).

In this paper, we are interested in another aspect that completes the span of the use of brackets which is the use of unnecessary, superfluous brackets. Sheppard (1916) very early suggested their use claiming that students "avoid the necessity of learning by rote the arbitrary and (to the young student) apparently meaningless rule that multiplications and divisions are to be performed before additions and subtractions" (p. 296). The issue has been examined sporadically and the findings are mixed. Some studies suggest the use of unnecessary brackets to emphasize the structure of the expression (Marchini & Papadopoulos, 2011). Others express their doubt about whether the use of unnecessary brackets benefits students (Gunnarsson et al., 2016). Marchini and Papadopoulos (2011) used unnecessary brackets in arithmetical expressions in grades 2 and 3 in Italy and Greece for identical arithmetical expressions

such as □ + 4 = 9 and (□ + 4) = 9 and found that their use considerably improved the number of correct answers by helping students visualise the content of these brackets as a single entity. On the other hand, Gunnarsson et al. (2016) transformed expressions of the form $a \pm b \times c$ to $a \pm (b \times c)$ to emphasize the priority of multiplication over addition or subtraction aiming to highlight the structure of the expressions thus facilitating the conventional rules. In the end, it was found that the use of unnecessary brackets did not enhance students' performance when learning the order of operations. All these studies share a common characteristic. The introduction of the unnecessary brackets was made by the researchers aiming to serve certain teaching or research aims.

In this landscape, our interests are focused on the use of unnecessary brackets that are initiated by the students themselves while evaluating algebraic expressions. So, our research question becomes: In what ways students' structure sense is displayed through the different uses of unnecessary brackets?

THE SETTING OF THE STUDY

The participants were 181 Grade 10 and 11 (16-17 years old) students from rural schools in Greece. According to the official curriculum, they had been taught the rules for the order of operations, exponents, and algebraic identities. The students were given a questionnaire with two groups of tasks. The first group included two tasks with exponents (Fig. 1), asking for the calculation of powers with positive and negative exponents when the base of the power is an integer or a rational number.

1) Calculate the value of the expression x^4 when:

a) $x = -2$ b) $x = \frac{1}{-2}$ c) $x = -\frac{2}{-3}$

2) Calculate the value of the expression x^{-2} when:

a) $x = -2$ b) $x = -\frac{1}{2}$ c) $x = \frac{2}{3}$ d) $x = \frac{2}{-5}$

Figure 1: First group of tasks

The second group included four items (Fig. 2) asking for substitutions in algebraic expressions.

3) Calculate the values of the following expression a^{-2b} when:

i) $a = -2$ and $b = -1$ ii) $a = -1$ and $b = -2$

4) Calculate the values of the following expression x^2y when $x = -3$ and $y = -1$

5) Calculate the values of the following expression $xy - x^2$ when $x = -2$ and $y = -3$

6) Verify that $(a - b)^2 = a^2 - 2ab + b^2$, by calculating separately the numerical result of the left- and right-hand side for each of the following cases:

i) $a = -2$ and $b = -3$ ii) $a = -2xy$ and $b = -xy$

Figure 2: Second group of tasks

In all items the use of brackets is imperative to get the correct result. Moreover, the items of the second group demand both algebraic manipulation and knowledge about integer exponents. The students worked individually for about 10-15 minutes and their written solutions constitute the data of this study. Data were analyzed in the context of qualitative content analysis (Mayring, 2014) by initially identifying instances of the use of unnecessary brackets. Then, these instances were further examined and categorized in terms of their connection to the issue of structure sense and the strategy the students followed.

RESULTS AND DISCUSSION

The analysis resulted in 203 instances of the use of unnecessary brackets distributed in three main categories. Students' writings illustrate the use of unnecessary brackets to show (i) the substitution of certain values (165 instances), (ii) to show the application of certain mathematical knowledge (7 instances), and (iii) to show the grouping of terms (31 instances).

Use of unnecessary brackets to show substitution

In this case, students used unnecessary brackets to show how they substituted the asked values (numbers or monomials) to the given variables. In Task 1, some students translated first the exponent x^4 as $x \cdot x \cdot x \cdot x$, then substituted the given number and used brackets to enclose each substitution. In each case, the first pair of brackets is unnecessary (Fig. 3, left). Similar use of brackets can be seen in the example drawn from Task 5 (Fig. 3, right).

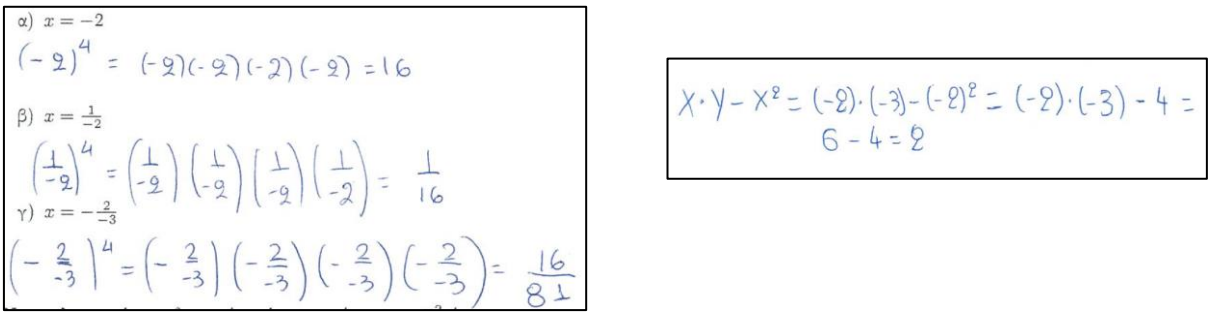


Figure 3: Students' scripts illustrating substitution of numerical values

The same happened in Task 6 when students substituted the variables a and b with the monomials $-2xy$ and $-xy$ respectively (Fig. 4). The student initially rewrote the expression and then when substituting the monomials in a and b they added brackets. In the substitution of the monomial $-2xy$ to a the brackets are not needed. However, the student kept using them to possibly illustrate the term a and its relation to the rest of the expression.

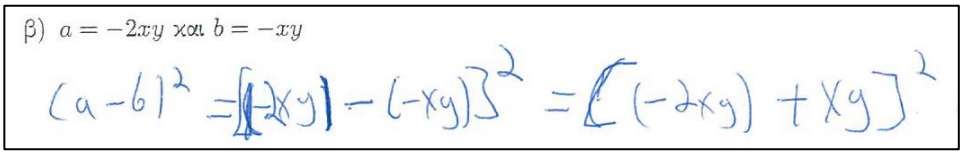


Figure 4: Substituting monomials

This use of unnecessary brackets is aligned with what Novotná and Hoch (2008) call "substitution principle", an important feature of structure sense referring to the replacement of a variable by a compound term (the product $-2xy$ in our case). In total 165 instances of using unnecessary brackets for substitution were identified in students' written answers.

Use of unnecessary brackets to show the application of certain mathematical knowledge

In this category, students used unnecessary brackets to show how they applied certain rules. These students seem to want to manipulate the expressions in line with the appropriate rules and explicitly highlight the structures within an expression linking to the rule. Seven such instances were identified in students' answers. In Figure 5, the student used the distributive law to calculate the result of the product $(-2+3)\cdot(-2+3)$. Parentheses and square brackets have been used to show the four terms that are formed when applying the distributive law. However, for three of the four resulting terms, the brackets are unnecessary.

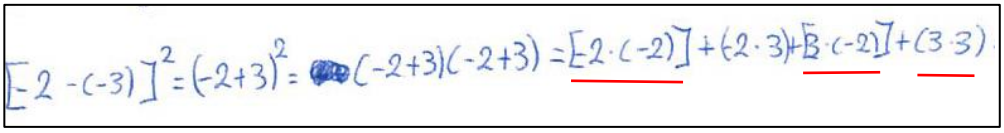


Figure 5: Applying the distributive law – Student's answer to 6i)

Similarly, in Figure 6, the student used unnecessary brackets to show how the following exponent rule $\left(\frac{a}{b}\right)^k = \frac{a^k}{b^k}$ was applied. So, she used unnecessary brackets to illustrate terms a and b (2 and 3 respectively) raised to the same exponent.

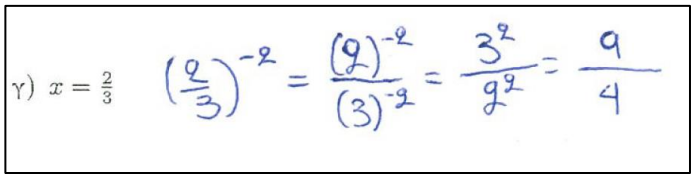


Figure 6: Applying exponential properties

Having a structure sense for the distributive law means that students easily recognize the forms $ab + ac$ and $a(b + c)$ as equivalent (Schüler-Meyer, 2017). Mastering the syntax or the rules for manipulating symbolic writings is closely connected with possessing structure sense (Linchevski & Livneh, 1999).

Use of unnecessary brackets for grouping purposes

Another category of using unnecessary brackets was the grouping of terms. The difficulty of the tendency to detach a numeral from the preceding minus sign in the grouping of numerical terms has been discussed in the literature (Linchevski & Livneh, 1999). Structure sense demands an understanding of how terms in arithmetic expressions are grouped and this grouping cannot be made arbitrarily (Gunnarsson &

Karlsson, 2015). In our study, we identified two different kinds of using unnecessary brackets for grouping: (i) to keep certain terms of the expression together, and (ii) to keep the result of the operation separate.

In the first case, students use unnecessary brackets to keep together either the whole expression or parts of an expression. In Figure 7 (left) square brackets have been used to keep together the whole right part of the identity $(a-b)^2$after substituting variables a and b with the monomials $-2xy$ and $-xy$ respectively. In the same figure on the right, the two examples show the use of unnecessary brackets to keep together a part of the expression. More precisely, the brackets are used to keep together the $2ab$ part of the identity when substituting a and b with numbers (above) or the ab part when substituting a and b with monomials (below).

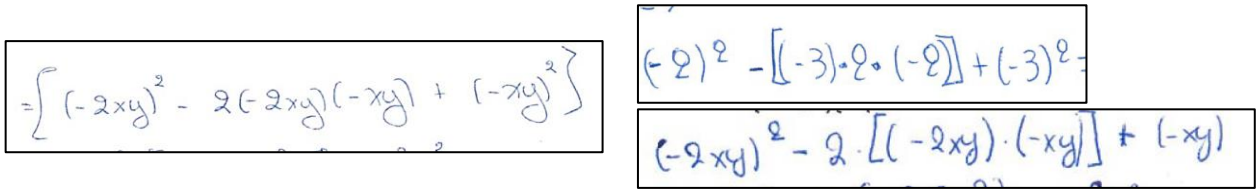


Figure 7: Grouping for keeping certain terms of the expression together

In the second case, the students, no matter whether they worked with numbers or variables, used unnecessary brackets to show the result of their calculations. In Figure 8 (left) unnecessary brackets have been used to show the result of $(-2)^2$ (above) and $(-2)\cdot(-3)$ (below). On the right side of Figure 8 unnecessary brackets are used to show the result of the evaluation of the algebraic expression.

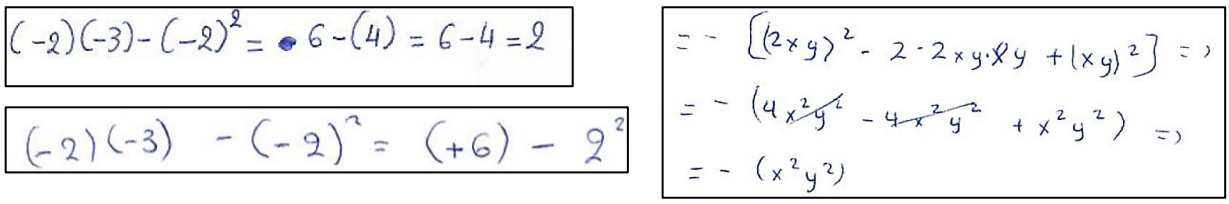


Figure 8: Grouping for keeping the result of the operation separate

In the examples above unnecessary brackets act as a perceptual grouping mechanism within a mathematical expression to support the successful accomplishment of the tasks. They form a common visual area inside the expression that draws the attention of the student to this specific part of the expression (see the work of Landy & Goldstone (2010) who examined analogous visual evidence in students' work that had been formed by leaving a noticeable blank space between some terms of the expression). This kind of grouping enables fast, effortless apprehension of the internal structure of expressions (Braithwaite et al., 2016). In total 31 instances of grouping were identified in the students' written answers.

CONCLUSIONS

In this study, students' use of unnecessary brackets while evaluating algebraic expressions is examined. The use of unnecessary brackets, together with the use of

necessary and mental brackets completes the span of the brackets' use in relation to structure sense. The connection between the last two uses of brackets (necessary and mental) with students' structure sense has already been highlighted (Linchevski & Livneh, 1999; Papadopoulos & Gunnarsson, 2020). The aim of this paper is to illustrate the connection between students' use of unnecessary brackets and structure sense.

It is worth mentioning that in contrast with the relatively small number of studies focusing on unnecessary brackets, in this study it is the students (instead of the researchers) who initiated the use of unnecessary brackets while evaluating algebraic expressions. So, the students intentionally changed the visual presentation of the expression, and these changes seem to serve as powerful cues that guide students' attention on specific parts of the expressions and support them to accomplish the evaluation of the expression. This visual enhancement of the structural elements of the expression constitutes one of the possible ways to support the development of students' structure sense (Marchini & Papadopoulos, 2011).

The detailed analysis of students' written responses in this study resulted in exemplifying three different uses of the unnecessary brackets. First, for substitution purposes, an important feature of structure sense is when brackets were used to show the replacement of a variable by a number or a compound term (Novotná & Hoch, 2008). Second, for showing the application of certain mathematical knowledge such as the exponential properties or the distributive law, thus exhibiting mastering the syntax or the rules for manipulating symbolic writing, which is indicative of structure sense (Schüler-Meyer, 2017; Linchevski & Livneh, 1999). Third, for grouping purposes, where unnecessary brackets were used as a perceptual mechanism that allowed them to apprehend the internal structure of given expressions (Braithwaite et al., 2016).

It can be said that the use of unnecessary brackets, along with other perceptual cues such as spacing (Landy & Goldstone, 2010) can be utilised as ways to support students' development of structure sense. If students are exposed very often to unnecessary brackets as a visual element to retain their attention to certain parts of an expression, then it might be possible for students to enhance their understanding of the structural elements of the expression.

It is important to note that our study's small number of participants does not allow to generalize the findings. However, they are encouraging enough to prompt a broader study including students across all educational levels and exploring the use of unnecessary brackets in different mathematical domains. In our study, we used tasks from one mathematical domain, namely algebra and focused on students' written solutions. Further investigation on students' use of unnecessary brackets in different mathematical domains (e.g., calculus) and additional data (e.g., students' interviews) would provide further insight in students' use of unnecessary brackets and their structure sense.

REFERENCES

Braithwaite, D. W., Goldstone, R. L., van der Maas, H. L. J., & Landy, D. H. (2016). Non-formal mechanisms in mathematical cognitive development: The case of arithmetic. *Cognition, 149*, 40–55.

Gunnarsson, R., & Karlsson, A. (2015). Brackets and the structure sense. In O. Helenius, A. Engström, T. Meaney, P. Nilsson, E. Norén, J. Sayers & M. Österholm (Eds.), *Proceedings of the 9th Swedish Mathematics Education Research Seminar* (pp. 47–55). MADIF.

Gunnarsson, R., Sönnerhed, W. W., & Hernell, B. (2016). Does it help to use mathematically superfluous brackets when teaching the rules for the order of operations? *Educational Studies in Mathematics, 92*(1), 91–105.

Hoch, M., & Dreyfus, T. (2004). Structure sense in high school algebra: The effect of brackets. In M. J. Høines & A. B. Fuglestad (Eds.), *Proc. 28th Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 3, pp. 49–56). PME.

Hoch, M., & Dreyfus, T. (2006). Structure sense versus manipulation skills: An unexpected result. In J. Novotná, H. Moraová, M. Krátká, & N. Stehlíková (Eds.), *Proc. 30th Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 3, pp. 305-312). PME.

Landy, D., & Goldstone, R. L. (2010). Proximity and precedence in arithmetic. *The Quarterly Journal of Experimental Psychology, 63*(10), 1953–1968.

Linchevski, L., & Livneh, D. (1999). Structure sense: The relationship between algebraic and numerical contexts. *Educational Studies in Mathematics, 40*(2), 173–196.

Marchini, C., & Papadopoulos, I. (2011). Are useless brackets useful for teaching? In B. Ubuz (Ed.), *Proc. 35th Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 3, pp.185–192). PME.

Mayring, P. (2014). Qualitative content analysis: Theoretical foundation, basic procedures and software solution. Beltz.

Novotná, J., & Hoch, M. (2008). How structure sense for algebraic expressions or equations is related to structure sense for abstract algebra. *Mathematics Education Research Journal, 20*(2), 93-104.

Papadopoulos, I., & Gunnarsson, R. (2020). Exploring the way rational expressions trigger the use of "mental" brackets by primary school students. *Educational Studies in Mathematics, 103*(2), 191-207.

Papadopoulos, I., & Thoma, A. (2022). Mental brackets and their use by high school students in arithmetic and algebra. *International Journal of Science and Mathematics Education* (in press).

Schüler-Meyer, A. (2017). Students' development of structure sense for the distributive law. *Educational Studies in Mathematics, 96*, 17–32.

Sheppard, W. F. (1916). The Use of Brackets in Arithmetic. *The Mathematical Gazette, 8*(124), 296–296.

TYPES AND FEATURES OF DIALOGICAL TASKS FROM MATHEMATICS TEACHERS' PERSPECTIVE

Reut Parasha and Boris Koichu

Weizmann Institute of Science, Israel

This study aims to examine the types of mathematical tasks that teachers select or modify to elicit argumentative dialogue among students, and to identify the dialogical characteristics attributed to these tasks by teachers. The data were collected in a task-based PD program aimed at fostering dialogical teaching in mathematics classrooms. The results indicate that when it comes to the independent selection or modification of a task, many types are considered, but most of the teachers prefer tasks that based on explication of students' mistakes and misconceptions. Additionally, we identified three features of tasks that make them dialogical from the teacher's perspective: Triggering cognitive conflict, mathematical complexity, and encouraging multiple solutions or approaches to the same problem.

INTRODUCTION

The importance of dialogical teaching, which encourages students to engage in dialogic argumentation during collaborative problem solving, is broadly recognized in the professional literature. Researchers (e.g., Swan, 2007; Sullivan et al., 2012) have designed various types of tasks to facilitate learning via reasoning and discussion in the mathematics classroom. To trigger discussions and support productive conversations, the literature recommends mathematical tasks to bear elements of disturbance (Zazkis & Mamolo 2018), surprise (Swan, 2007), cognitive conflict (ibid), complexity (Sullivan et al., 2012), multiple solutions and approaches (Leikin, 2011), open questions and prompts for deliberating reasoning (Sullivan et al., 2012), uncertainty (Schwarz & Baker, 2016), and negotiability in that students can elaborate on different perspectives based on their previous experience and knowledge (ibid).

However, the teacher perspectives on feasibility of and opportunities for stimulating argumentative dialogue among students in a mathematics classroom through task selection/modification is still insufficiently captured (Kaur & Chin, 2022). To address this gap, this study examines types of tasks that teachers consider as dialogue-eliciting, and strives to identify the main features they attribute to these tasks to encourage dialogue among students. Our research questions are as follows. While participating in a task-based professional development (PD) program aimed at fostering productive dialogue in the mathematics classroom, (1) What types of mathematical tasks do teachers independently select/modify in order to trigger argumentative dialogue among their students? (2) From the teacher's perspective, how might these tasks promote argumentative dialogue among students in a mathematics classroom?

2023. 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. 19-34). PME 46.

THEORETICAL BACKGROUND

Schwarz and Baker (2016) characterize argumentation that is desirable in education as a kind of dialogue that integrates rigorous reasoning and respectful reference to the other. They show that dialectical moves, as opposed to consensual explanations, predict conceptual gains and suggest that this type of talk is most productive for learning. To this end, we consider *argumentative dialogue* as a dialogue in which students collaboratively explore differing views, and *dialogical tasks* as mathematical tasks that have the potential to elicit respectful disagreement as students work together.

Despite the literature's broad acknowledgment of the importance of dialogical teaching, and despite teachers' apparent desire to develop their students' dialogical abilities, systematic enactment of dialogical tasks is not yet widespread (Schwarz & Baker, 2016).

Choy and Dindyal (2021) note that although most teachers use "standard" (i.e., procedural, textbook-like) tasks available to them, they are able to modify them so as to provoke meaningful discussion among students, provided that the teacher recognizes the affordances inherent in the task. Crespo and Harper (2020) suggest that teachers' exposure to and experience with a variety of structures and task types can help them in selecting tasks or modifying standard tasks. Jones and Pepin (2016) also point out the importance of teachers being active participants in selecting, adapting, or designing tasks in order to develop their agency over the tasks and over their enactment.

In light of the above, we adopted the following key principles regarding the selection or development of dialogical tasks by and with teachers: 1. A standard task can be transformed into a dialogical task by modifying it. 2. With appropriate support, teachers are able to modify the tasks available to them to make them more dialogical. 3. Task-based PD that aims to promote productive dialogue in the mathematics classroom should provide teachers with a repertoire of experiences with a variety of dialogical types of tasks and encourage them to independently select tasks that have dialogical spirit or to modify standard tasks into dialogical tasks. Types of dialogical tasks that can be used in such a PD include open-ended tasks (Zaslavski, 1995); who-is-right tasks (Koichu et al., 2021); evaluating mathematical statements tasks (Swan, 2007); sorting tasks (ibid); problem-posing tasks (ibid) and more.

METHOD

Context, participants and data sources

The data were collected in 2018-2021, as part of the activity of DIALOGOS project. The goal of the project was to develop dialogical instructional materials and facilitate dialogical teaching in science, mathematics and philosophy in six secondary schools in the central part of Israel. The mathematics R&D group of the project collaborated with 26 in-service teachers (4 males and 22 females) who had teaching experience ranging from one to 20 years. The teachers, who as a rule did not have prior experience in dialogical teaching, participated in a PD course, which was conducted by the

DIALOGOS team. In the mathematical part of the PD (20 to 30 hours per year), the main effort was devoted to experiencing, designing and reflecting on dialogical mathematical tasks in accordance with the key principles described in the previous section.

The DIALOGOS team collected data from various sources, two of which are relevant to the concerns of the study reported in this paper: written final assignments of the PD, and video-recorded oral presentations during the PD meetings. The final assignment was given in two formats for the teachers to choose from. The first was to plan a lesson that would encourage argumentative dialogue among students. The second format required selecting two "standard" tasks and modifying them to evoke argumentative dialogue among students. The instructions for the oral presentation were to present a mathematical task that evoked an interesting dialogue in a class, or a task having the potential to do so. In each of the options, the teachers had to explain why the task they proposed was expected to stimulate the desired dialogue. Some teachers submitted the assignment in pairs, while others submitted them individually. In terms of the data analysis and for answering the research questions at stake there was no significant difference between teacher products following different formats of the final assignment and the presentation. In summary, 60 dialogical tasks chosen or designed by 26 teachers were analysed, along with their considerations of why these tasks would evoke argumentative dialogue among secondary school students.

DATA ANALYSIS

The data were analysed in two steps correspondingly to the research questions (RQs): identification of the type of the task (i.e., what the learner is asked to do), and of the teacher considerations regarding the dialogical potential. To classify the tasks, we used deductive coding methods (Creswell, 2014) by coding the tasks in terms of categories known from the literature (e.g., a sorting task, a problem-posing task, etc.). We also employed inductive analysis methods (Creswell, 2014) to categorize tasks that have not been described in the literature. For the second step, we used an open, iterative coding process (ibid) in order to create categories that reflect the main characteristics of a dialogical task from the teacher’s perspective. Two guiding questions accompanied this step in the analysis: why did the teacher consider the task to be dialogue-eliciting? In the teacher’s view, what would the students disagree about? The main codes used to classify the tasks are presented below.

Monitoring & Evaluation tasks: Tasks requiring evaluating, comparing, or monitoring given mathematical statements or solutions. These tasks encourage students to evaluate and compare alternative reasoning approaches (Swan, 2007). For example, tasks requiring to decide which of the given solutions is correct (i.e., who-is-right tasks), tasks requiring to identify and correct mistakes in reasoning, or evaluate the correctness of several mathematical statements. Two sub-categories emerged in this category: Analyzing reasoning and solutions (ARS) and Evaluating mathematical statements (EMS).

Particularly challenging tasks: Tasks requiring some creative effort and higher-level thinking to solve instead of a direct application of a procedure (Yeo, 2017). The category consists of two sub-codes: Problem-solving task (PST), and Qualitative reasoning task (QRT).

Investigation tasks: Tasks requiring investigation and discovering of underlying patterns or structures of mathematical objects or situations encountered in everyday life. (This is in contrast to tasks that require investigation to solve a problem, Yeo, 2017). Three sub-codes emerged from the data in this category: Open investigation tasks (OIT), Closed investigation tasks (CIT) and Noticing tasks (NOT).

Multiplicity-eliciting tasks: Tasks that include specific instructions to provide more than one method for solving a problem (MPM), to present multiple possible solutions (MPS), or express opinions about multiple methods/solutions (e.g., "what is the best method?") (OPI).

Classifying mathematical objects (CMO): Tasks requiring to examine objects, and classify or sort them according to their different attributes (Swan, 2007).

Problem-Posing (PP): Tasks requiring students to devise their own mathematical problems or questions (Swan, 2007).

Procedure application tasks (PAT): Tasks calling for direct application of procedures (Yeo, 2017), which are generally perceived to be of low cognitive demand and can be solved by students in relatively short time. The format of PAT is usually "solve", "calculate", "prove that...", or "find".

Other - Tasks not covered by any of the above codes.

FINDINGS

To answer RQ1, we report in Table 1 on the types of tasks the teachers suggested as dialogical tasks. Some of the teachers presented more than one task, and some of the tasks were coded with more than one code. Accordingly, percentages in Table 1 sum up to more than 100%. Of the 60 dialogical tasks, Monitoring & Evaluation tasks were the most common, presented by 77% (out of 26) of the teachers. The other task types presented in the PD sessions, such as Classifying of mathematical objects, Problem-posing, and Open-ended tasks, were relatively rare or did not appear at all. We deem this finding quite surprising because during the PD many teachers reported that they never use tasks belonging to the Monitoring & Evaluation category in their lessons. One of the teachers even referred to who-is-right tasks as "childish" in one of the PD meetings, but eventually she chose to present this type of tasks in her final essay. It is also interesting to see that 54% of the teachers chose Procedure application tasks (PATs) as dialogue-eliciting tasks. Indeed, why might a simple, procedural task trigger argumentative dialogue among students? A close look at the data reveals several answers to this question. First, 14 out of 24 PATs were augmented by the teachers with additional types of tasks. Second, sometimes PATs were used as a prerequisite for the dialogue foreseen. In additional cases, a PAT on its own was considered as evoking

argumentative dialogue because of the expected procedural difficulties or common mistakes that might occur in the classroom.

Table 1: Types of tasks and their frequency of appearance

Code (type of task)	Number (%) of tasks N=60	Number (%) of teachers N=26
Monitoring & Evaluation tasks	27 (45%)	20 (77%)
Particular Challenging tasks	10 (17%)	12 (46%)
Investigation tasks	14 (23%)	11 (42%)
Multiplicity-eliciting tasks	10 (17%)	7 (27%)
Classifying mathematical objects	2 (3%)	2 (8%)
Problem-Posing	1 (2%)	1 (4%)
Procedure application tasks	24 (40%)	14 (54%)
Other	4 (7%)	4 (15%)

As for RQ2, three (not mutually exclusive) features of dialogical tasks were inductively distilled from the teachers' explanations.

Cognitive conflict – a feature of a task that causes cognitive dissonance in learners, as a result of the exposure to conflicting conceptions, various mistakes, or inconsistent ideas that emerge during the task solution.

The dialogue that may be triggered by the "cognitive conflict" includes expressions of disagreement, counter-arguments and attempts at mutual persuasion. At the same time, it can include dialogical moves aimed at reflecting on the dissonance revealed. From the teacher's perspective, this feature contributes aspects of disturbance and surprise to the task-solving experience, as shown in the following examples taken from final essays as particularly illustrative.

Teacher#10: "An optimal mathematical dialogue can be created by presenting the students' errors and a discussion following the errors."..." I presented them with incorrectly solved exercises in order to cause a cognitive conflict" (Final essay, lesson plan, ARS tasks).

Teacher#16: "A dialogic discourse that confronts the students with mistakes they made in the solution... when there were mistakes, and students explained to each other, and insisted on having an explanation of why their way of thinking was incorrect, like "I understood what you explained, but wait, why is what I'm saying wrong?" That is, they were not satisfied with the correct answer but really wanted to understand the errors in their thinking" (Final essay, reflecting on ARS+QRT tasks that evoked dialogue in class)

Teacher#11: "At this point, a dialogue and a small "storm" emerges in the class, because how could a result different from mine being obtained?" (Final essay, a planned dialogic lesson with PAT+ARS tasks).

Mathematical Complexity – a feature of a task to be cognitively or mathematically demanding for the students. According to the teachers, the mathematical complexity of a task can be manifested in two ways: i) **Procedural complexity** – the task involves

technical mathematical procedures that are difficult, long, or not readily accessible to the solvers though still within their reach. ii) **Problem-solving complexity** – the task involves the need to invent a previously untaught solution path. Namely, the solver is unacquainted with some of the content required, or with the solution method, hence she does not have a readily accessible pathway for solving the question.

In this case, a dialogue can take the form of students sharing knowledge and ideas, or of disagreements over the correctness of proposed solution strategies and approaches. The complexity element creates a sense of uncertainty among students in the task-solving experience and reinforces the need for collaboration and dialogue.

Teacher#15: "The source of the dialogue: How do you find the limits? How do you calculate the integral of this function?" (Final essay, reflecting on a PAT that evoked dialogue in class through procedural complexity).

Teacher#1: "The dialogical tasks will stimulate brainstorming, thinking and discussion as they are more complex and deeper than the original questions, and their answers are not trivial" (Final essay, modifying PAT into CIT).

Teacher#20: "...There are constraints, and you have to reach the largest volume...the students: "No, I have a better idea! Let's try this way, no!" This is how they try to convince each other." (Oral presentation, reflecting on a PST that has evoked a dialogue in the class).

Multiplicity – a feature of a task that explicitly elicits different (correct) solutions or approaches. The "Multiplicity" of the task can trigger a dialogue because when students take different approaches to the same task, they are forced to convince others of the merit of their approach. The multiplicity feature also provides different points of access to students in heterogeneous groups, allowing more students to actively participate in the discussion.

Teacher#21: "The task is an opportunity for the children to discuss their different ways of solving, even though at the end there is only one answer." (Final essay, lesson plan, MPM task).

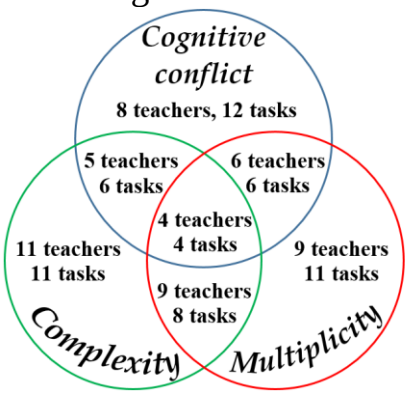
Teacher#24: "Question No. 2 [of the task designed] allows multiple correct answers... the significant dialogue is expected in question 2... around the different ways by which it is possible to prove that the square is a rectangle" (Final essay, lesson plan, OIT).

Teacher#10: "Open-ended tasks have the potential to stimulate discussion among the students, because they have multiple solutions, therefore all student in the class have the opportunity to find different solutions according to their level of knowledge and in a class discussion to receive many correct solutions." (Final essay, lesson plan, OIT)

Figure 1 shows the distribution of the teachers and the number of tasks related to seven different characterizations of dialogic tasks resulting from the three features. To recall, the same task could be characterized by more than one feature. For example, a task with specific instructions to solve a problem in multiple methods (MPM) was sometimes considered dialogical by teachers because of its Multiplicity as well as its

Complexity. As one can see, most of the teachers characterized their tasks using one or two of the features, and only 4 teachers attributed to their tasks all three features.

Figure 1 – Seven features of dialogical tasks based on three inductive themes



CONCLUDIND REMARKS

Despite the extensive literature on dialogical tasks and their desirable characteristics, there is a lack of knowledge about the tasks teachers perceive as dialogical. In our study (RQ1), we found seven main types of tasks the teachers see as (potentially) dialogical, and reported their frequencies of appearance in the data pool (Table 1).

For us, three findings are particularly notable. First, of the various types of dialogical tasks presented in the DIALOGOS PD sessions, most teachers presented tasks that require monitoring and evaluation based on explicating students' errors and misconceptions. Second, the other types of tasks presented in the PD sessions, such as classifying objects, problem-posing, and open-ended tasks, were rare in teachers' essays. Instead, several types of closed investigation tasks, which teachers described as stimulating dialogue because of their complexity, were chosen, though these were not presented or discussed in the PD meetings. Third, the teachers proposed as dialogical a relatively large number of unmodified PATs, while recognizing the potential of these tasks for dialogicity in their pedagogical moves rather than in the task formulations. For example, such tasks could be used as a reference for building arguments for the forthcoming "non-standard" tasks. These findings can be of importance for understanding which types of dialogical tasks teachers are likely to use and which practices they are comfortable with, such as identifying and discussing student errors. The findings also empirically support Choy and Dindyal's (2021) idea of noticing the (dialogic) affordances of a (standard) task by anticipating possible difficulties that students may face.

With respect to RQ2, we identified three features of tasks that teachers see as potentially dialogical: cognitive conflict, mathematical complexity, and multiplicity. On the one hand, these three features are well established in the literature. On the other hand, their appearance in our data pool adds ecological validity to the literature and encapsulates it into three operational characteristics that, as we know now, are relevant for teachers. Accordingly, we deem these findings of potential importance for promoting argumentative dialogue in the classroom, as follows: when teachers

recognize the dialogic affordances inherent in the task, they are more likely to be able to harness them to promote argumentative dialogue.

Hopefully, the three features of dialogical tasks identified in our study from the teachers' perspective can be used in practice in two complementary ways: as a seed of a framework for characterizing dialogical situations as planned and enacted, and as a pedagogical tool for teachers to select, design and enact dialogical tasks.

Acknowledgements The study is partially supported by the Israeli Scientific Foundation, Grant No. 2699-17. We thank Jason Cooper for his valuable suggestions.

REFERENCES

Choy, B. H., & Dindyal, J. (2021). Productive teacher noticing and affordances of typical problems. *ZDM–Mathematics Education*, 53(1), 195-213.

Crespo, S., & Harper, F. K. (2020). Learning to pose collaborative mathematics problems with secondary prospective teachers. *International Journal of Educational Research*, 102.

Creswell, J. W. (2014). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th edition). Upper Saddle River: NJ: Prentice Hall.

Jones, K., & Pepin, B. (2016). Research on mathematics teachers as partners in task design. *Journal of Mathematics Teacher Education*, 19(2-3), 105-121.

Kaur, B., & Chin, S. L. (2022). Nature of mathematics tasks and what teachers do. *Current Opinion in Behavioral Sciences*, 46, 101169.

Koichu, B., Parasha, R., & Tabach, M. (2021). Who-Is-Right tasks as a means for supporting collective looking-back practices. *ZDM–Mathematics Education*, 53(4), 831-846.

Schwarz, B. B., & Baker, M. J. (2016). *Dialogue, argumentation and education: History, theory and practice*. Cambridge University Press.

Sullivan, P., Clarke, D., & Clarke, B. (2012). *Teaching with tasks for effective mathematics learning* (Vol. 9). Springer Science & Business Media.

Swan, M. (2007). The impact of task-based professional development on teachers' practices and beliefs: A design research study. *Journal of Mathematics Teacher Education*, 10(4-6), 217-237.

Yeo, J. B. (2017). Development of a framework to characterise the openness of mathematical tasks. *International Journal of Science and Mathematics Education*, 15(1), 175-191.

Zaslavsky, O. (1995). Open-ended tasks as a trigger for mathematics teachers' professional development. *For the Learning of Mathematics*, 15(3), 15-20.

Zazkis, R., & Mamolo, A. (2018). From disturbance to task design, or a story of a rectangular lake. *Journal of Mathematics Teacher Education*, 21(5), 501-516.

ARE EXPERTS' NOTICING FOCUSES REGARDING THE LEARNING POTENTIAL OF TASKS AND ITS USE CONSISTENT ACROSS INSTRUCTIONAL SITUATIONS? A SECONDARY ANALYSIS

Josephine F. Paul[1], Anika Dreher[2], Ting-Ying Wang[3], Feng-Jui Hsieh[3], Linn Hansen[2], Anke Lindmeier[1]

[1]University of Jena, Germany, [2]Freiburg University of Education, Germany, [3]NTNU Taipei, Taiwan

Professional noticing is an important aspect of teacher expertise. Research has shown that noticing might vary between cultures. Prior studies indicate that culture-specific norms influence noticing. But, these findings stem from group comparisons. The reported study aims at understanding how individual focuses may influence noticing in addition to cultural influences. It uses noticing vignettes to investigate the aspects Western (German, N=17) and East Asian (Taiwanese, N=19) professors in mathematics education (experts) focus their attention on when evaluating the potential of tasks for mathematical learning and its use in instruction. We present preliminary results regarding the consistency of the focuses across situations, identify five types of experts regarding their focuses, and discuss implications for further noticing research.

THEORETICAL BACKGROUND

Noticing, beliefs, and knowledge

We follow the idea of Sherin (2017) and understand noticing as professional processes that include *selective attention* to specific "aspects of classroom situations that are relevant for instructional quality" (Dreher et al., 2021, p. 90) and *knowledge-based reasoning* of the noticed instructional aspects. We describe selective attention as the specific *focus* persons have during noticing. We understand selective attention and knowledge-based reasoning as cyclical interacting subprocesses (Dreher et al., 2021). Just as the focused aspects determine how one interprets what is noticed, the professional knowledge and beliefs one draws on determine the focus. Beliefs are "understandings, premises, or propositions about the world that are thought to be true. […] Beliefs might be thought of as lenses that affect one's view of some aspect of the world" (Philipp, 2007). Thus, if different persons focus their attention on different aspects, this may be due to differences in beliefs or professional knowledge.

To investigate noticing, it is common to use video, comic, or written vignettes. Written vignettes benefit from an easy adaption during the development process and allow for a representation of an instructional situation in a simplified manner (Dreher et al., 2021). Teachers typically have to analyse the instructional situations, and the responses to vignettes can be used to (1) identify the teachers noticing focus and (2) infer what knowledge guided their reasoning. In this report, we understand the *knowledge* somebody draws on when noticing as an overarching term for professional knowledge

2023. 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. 27-34). PME 46.

(CK, PCK, PK) *and* beliefs. Thus, we also understand evaluations based on beliefs as a knowledge-based interpretation within the noticing framework.

Beliefs can refer to the facets of professional knowledge. They influence building and using professional knowledge and vice versa. Similar to culture-specific norms, beliefs are acquired through "enculturation and social construction" (Pajares, 1992, p. 316). Furthermore, people form so-called belief systems (Pajares, 1992), which are considered as the entirety of beliefs a person holds. These systems are individual and do not require a general agreement or an agreement reached in groups (Nespor, 1987), as norms do. Likewise, they do not have to be consistent, and even if people share some of the same beliefs (e.g., due to underlying culture-specific norms), they can be very differently integrated into their belief systems.

Word problems as specific tasks in Western and East Asian cultures

In a primary analysis within the "TaiGer Noticing" project, cultural differences regarding task potential and its use were investigated (Lindmeier et al., 2022). As large disparities within the German and within the Taiwanese expert samples were discovered, we draw attention to the individual level. Therefore, aspects different persons might focus on when evaluating task potential for mathematical learning and its use should be identified by considering the previous research referring to mathematical tasks and their role in instruction. Mainly tasks provide opportunities to foster students' mathematical thinking and learning as well as students' mathematical engagement (e.g., Stein & Lane, 1996). Word problems as a specific type of task are defined as verbal problem descriptions, including questions students should answer by applying mathematical operations (Verschaffel et al., 2020). They are considered to have a high potential for students' mathematical learning if they are aligned with the learning objectives, encourage students to argue mathematically, or allow multiple solutions. However, whether a task's potential is used appropriately depends on its implementation by the teacher (e.g., Stein & Lane, 1996).

Beyond such general descriptions of potentials of word problems, cultural differences may also be relevant. In Western countries, word problems are considered to have high potential when they refer to real situations, promote the creation of mathematical models (process orientation), and foster individual learning (Leung, 2001). Whereas in East Asian countries, a task is considered to hold high potential when it promotes focusing on correct calculations and results (product orientation), as well as fosters learning, practising, and repeating strategies (subject content) (e.g., Leung, 2001).

RESEARCH QUESTIONS

Against this background, we ask: (Q1) Which noticing focuses do experts have when evaluating task potential and its use? (Q2) Are the aspects that experts focus on the same across various instructional situations so that types of experts can be identified?

CONTEXTS AND METHODS OF THE STUDY

This study is part of the binational research project "TaiGer Noticing". The project investigates differences in culture-specific norms regarding teaching quality. In the first project phase, the noticing of experts from both countries was elicited. Group comparisons indicate that experts from Germany and Taiwan evaluate the potential of tasks and its use differently and in line with cultural expectations (Lindmeier et al., 2022). This article reports a secondary analysis of the experts' responses following a different analysis approach. We examine how individual factors may influence noticing to better understand the interplay of different aspects in noticing processes.

Instruments

The instruments were designed and validated through a joint development approach of the German and Taiwanese research teams. The vignettes were translated and checked for ecological validity (i.e., instructional situations could have occurred in either country's secondary mathematics instruction). For further information about the development process, see Dreher et al. (2021). Each vignette includes a short description of an instructional situation, a picture of the task (Figure 1), and a fictitious transcript of a classroom interaction.

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 backboard.]
T: So, 52.5 ml is indeed the appropriate dosage for Paul. Well done, everyone.

Figure 1: Vignette Task1 (description, task, transcript)

The tasks were considered to have a high potential for mathematical learning from the perspective of the authoring team from one country but not necessarily the other. Primarily, the vignettes were designed to investigate culture-specific norms. Hence, each includes an anticipated breach of a norm (i.e., illustrates a non-optimal use of task

potential; for further information, see Dreher et al., 2021; Lindmeier et al., 2022). This study refers to three vignettes, but only one of them will be described in more detail in this report. For the other two, see Lindmeier et al. (2022). Despite displaying different instructional situations, they have a common frame of reference (mathematics instruction, same grade, word problems related to real-life situations), allowing for comparisons across the situations.

Vignette Task1 was designed in Germany. The vignette deals with proportional relationships. A text and an illustrative image show a real-life situation regarding medicine dosages. A specific dosage for a kid, Paul, shall be determined. The task requires finding at least two different solution strategies. Hence, the German authors assumed that the task potential lies in supporting students' flexible solving of real-life applications. Accordingly, the authors expect teachers to use this task to treat different solution strategies and highlight connections between them. In contrast, the teacher in the vignette just collects the solutions of two students verbally without further elaboration and only writes down a third solution using an equation on the blackboard.

Sample and procedures

The study participants were mathematics education professors in Germany and Taiwan. They were active in mathematics education research and teacher education and training. Assuming a participation rate of 50% and aiming for a sample of 15 experts in each country, a random sample of 30 professors in Germany was contacted. In Taiwan, only 32 professors met the criteria. All of them were contacted. In total, 36 experts participated: 19 Taiwanese professors (6 female, 13 male) from 10 universities and 17 German professors (7 female, 10 male) from 13 universities.

The experts were asked to respond with a short text to the question: "Please evaluate the teacher's use of the task in this situation and give reasons for your answer." The vignettes were presented in the experts' native language. The experts responded in German resp. Chinese writing. Hence, the response had to be translated into English as the common language within the research team. Afterwards, the experts' responses were analysed by qualitative content analysis to answer Q1. For this secondary analysis, a coding system of 13 codes regarding the focused aspects in the experts' responses was generated inductively (Table 1). The responses were analysed sentence by sentence, and a code was assigned if (positive and/or negative) annotations to the aspect occurred. Several codes could be assigned to each response. Codes were only assigned once per response, even if experts mentioned an aspect several times, for instance, at the end and the beginning of their response. It must be noted that this coding approach considerably differed from the approach in the original analysis, where the goal was to compare the noticing of the experts from the different cultures as groups.

After coding each response, we applied a person-related perspective to the analysis. To examine how consistently one expert focused on each aspect, it was counted in how many of their three responses (one to each vignette) an expert mentioned each of the 13 aspects. That gave us a dataset of 36 x 13 values from 0 (mentioned in none of the responses) to 3 (mentioned in each response). We applied a data-driven approach to

answer Q2 with this dataset: A hierarchical cluster analysis based on the code variables was conducted to determine different types of experts. An elbow plot was used to inform about the number of clusters. The types of experts were interpreted in terms of their characteristic patterns of focuses.

RESULTS

To answer RQ1, the focuses reflected in the experts' responses are presented in Table 1. As explained, a code was given when the aspect was mentioned in a positive or negative manner. For example, code 5 was given both when an expert said that the explanation and discussion of the solution were appropriate and when s/he noted that the solution should have been explored in more detail. It was possible to apply all codes across the three vignettes.

Code	Mentioning/evaluating (the)	Experts evaluate (the)	Number of code assignments in vignette Task1/2/4
V1	mathematical correctness	whether the presented solutions were correct	5/7/5
V2	mathematical completeness	solution as complete or mention missing mathematical aspects	3/3/5
V3	other mathematical aspects	use of terms, mentioning mathematical properties and solution strategies	6/3/8
V4	different solution paths	links between/fitting of/pros & cons of solution paths	27/21/32
V5	depth of explanation	explanation and discussion of the solution	17/22/13
V6	modelling aspects	link to the real-life situation or algebraization	3/25/7
V7	the value of the task	whether the benefit of the task is clear to the students	5/7/4
V8	students' understanding	whether the solution is comprehensible for all students	9/14/17
V9	Visualization/result consolidation	the need for a visualization or consolidation of the solution	11/5/6
V10	student activity	student's participation/using student's ideas	13/24/22
V11	teacher activity	teacher's activity positive/too directive; mention teacher's tasks	26/28/31
V12	appreciation of student activity	whether the teacher gave feedback/appreciated student's ideas	10/6/1
V13	pursuit of the learning objective	whether the teacher focused on his/her learning objective	7/3/4

Table 1: Coding system consisting of the identified focuses across all vignettes

The subsequent cluster analysis based on these variables revealed five clusters of experts (short: types). The types can be described as experts who

T1. have no specific focus; never focus on mathematical aspects (V1, V2, V3), V7 and V13 (2 Taiwanese, 4 German experts)
T2. mainly focus on V4 and V11 (5 Taiwanese, 6 German experts)
T3. mainly focus on V1, V5, V8, V9, V10, and V11; V4 is mentioned in each of the 3 responses (3 German experts)
T4. mainly focus on V4, V5, V10, and V11; V6 and V8 are at least mentioned in one response (4 Taiwanese, 2 German experts)
T5. mainly focus on V4, V10, and V11; V2 is never mentioned (8 Taiwanese, 2 German experts).

Note that *mainly focus* means that, on average, the experts assigned to the corresponding type mentioned an aspect on average in more than two of the three responses.

The T1-expert's responses shown in Figure 2 indicate no consistent focus. The aspects the expert refers to change between the different instructional situations, and none of the focuses is consistent across all three vignettes. In his or her response to Task2 s/he places the entire attention on the modelling aspects, whereas in responding to Task1 and 4 also, the different solution paths (V4) and the missing depth in elaborating the paths and the situation (V5) are mentioned.

Vignette	Response (partially abbreviated)	Codes
Task1	The teacher wants to promote flexibility in solving [real-life applications]. S/He does this by not only asking for results, but by encouraging and positively evaluating the variety of possible solution paths. […] However, some of the solution paths remain somewhat under-determined as for example "looking at the numbers". Only at the very end the actual result is recorded […]	V4, V5, V9, V12
Task2	The […] modelling task is largely ignored here and only the formal core of the task is discussed. Why it is necessary to determine the zero point in particular, i.e. why it is thus possible to mathematize in this way, is not discussed here, nor are the educated guesses in a) Thus the task is reduced to a formal calculation task.	V6
Task4	The task serves as an example for showing how to algebraize. Alternative algebraizations are taken up and are briefly addressed. It would be great to relate the two solutions to each other. The richness of the situation would be exploited even more thereby.	V4, V5, V6

Figure 2: Type T1 expert's responses to vignette Task1, 2 and 4. Note that the last column indicates all codes assigned to the full responses.

The consistent focuses in the T3-expert's response shown in Figure 3 are V4, V5, V10 and V11. S/He criticizes in each response that the teacher should have further explored the students' solutions and mentions dealing with different solution paths. Also, s/he draws attention to the improvement of the students' and the teacher's activity in each response. In addition, other aspects are mentioned once or twice, for instance, the modelling aspects (Task2) and the students' understanding (Task2, Task4).

Vignette	Response (partially abbreviated)	Codes
Task1	[…] when dealing with the students' responses, he/she confirmed the students' thoughts too quickly. […] [the teacher] did not conduct a discussion or explanation of how the students used these methods to obtain answers […] The same problem occurred to the reaction toward S3's proposing a proportional relationship. The teacher directly determined that the student used the function that relates the kg to the ml, and directly explained the idea for S3. […]	V4, V5, V10, V11
Task2	[...] it is a pity that the teacher did not allow the students to explain why they made these guesses […] When S1 stated […] that you just have to set f (x) = 0 and solve for x, the teacher immediately confirmed [the idea] and asked the students to solve the problem. This would perplex the students who did not yet understand that the problem could be solved by this method. The teacher should have asked why to use this method [and] could ask S4 to explain why -4 did not work [...]	V1, V4, V5, V6, V8, V10, V11, V13
Task4	[...] the teacher presented two approaches of setting equations and asked the students' preference […]. However, the students had different opinions, the teacher […] directly made a conclusion […] the students may still have difficulty to grasp the better way […] The teacher should have let the students make more tries and then compare, to gain their approaches of problem-solving.	V4, V5, V8, V10, V11

Figure 3: Type T3 expert's responses to vignette Task1, 2 and 4. Note that the last column indicates all codes assigned to the full responses.

DISCUSSION AND CONCLUSION

This study investigated whether experts' noticing focuses are consistent across different instructional situations regarding task potential and its use represented in written vignettes. We could identify 13 different focuses which occurred in the experts' responses. These noticing focuses (Table 1) align with the findings from research on mathematical tasks; for instance, tasks might have a high potential for students' mathematical learning when they align with the learning objectives (V13) or allow multiple solutions (V4). Looking more precisely at the experts' responses, it is evident that they are based on professional knowledge *and* beliefs. Exemplary, a belief that can be identified in the T3-expert's response to Task4 (Figure 3) is that *students do not understand or have difficulties solving tasks when they are not allowed to practice/try on their own.*

The results of the secondary analysis of a dataset of experts' noticing indicate that individual focuses may indeed prevail when noticing. Five types of experts can be differentiated regarding the focused aspects during noticing. Four types can be interpreted as describing consistent, yet different focuses: Persons of each type have in common that their main focuses across all vignettes are similar (T2-5). However, some experts in our sample also applied different focuses to different instructional situations presented in the vignettes (T1). For example, the vignette Task2 seems to include salient attractors regarding mathematical modelling, as the focus of some experts shifts to this aspect (see Task2 in Figure 2).

The study has several limitations. We will check whether the rating is intersubjectively applicable (intercoder reliability). The dataset may be considered marginally sufficient for the statistical analysis considering the number of persons in relation to the focused aspects. However, the resulting types of experts provide insight into the consistency of noticing focuses across situations. This short contribution could not report a more detailed presentation of each type and a comparison between German and Taiwanese experts. Also, cultural contrasts between German and Taiwanese experts' noticing were not discussed in detail in this contribution. A follow-up study will examine the connections between individual focuses, as reported in this paper, and the noticing of anticipated breaches of a norm as evaluated in the primary analysis of the "TaiGer Noticing" project. It will investigate how the different methodological approaches help us to understand how noticing processes depend on individual and cultural factors.

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 dealing with students' mathematical thinking. *ZDM*, *53*(1), 165–179. https://doi.org/10.1007/s11858-020-01197-z

Leung, F. K. S. (2001). In Search of an East Asian Identity in Mathematics Education. *Educational Studies in Mathematics*, *47*, 35–51.

Lindmeier, A., Wang, T.-Y., Hsieh, F.-J., & Dreher, A. (2022). The Potential of Tasks for Mathematical Learning and its Use in Instruction – Perspectives of Experts from Germany and Taiwan. *Proceedings of the 45th Conference of the IG PME*, *3*, 139–146.

Nespor, J. (1987). *The role of beliefs in the practice of teaching*. *Journal of Curriculum Studies*, *19*(4), 317–328.

Pajares, M. F. (1992). Teachers' Beliefs and Educational Research: Cleaning Up a Messy Construct. *Review of Educational Research*, *62*(3), 307–332.

Philipp, R. A. (2007). Mathematics teachers' beliefs and affect. In F. K. Lester (Hrsg.), *Second handbook of research on mathematics teaching and learning* (S. 257–315). National Council of Teachers of Mathematics.

Sherin, M. G. (2017). Exploring the Boundaries of Teacher Noticing: Commentary. In E. O. Schack, M. H. Fisher, & J. A. Wilhelm (Hrsg.), *Teacher Noticing: Bridging and Broadening Perspectives, Contexts, and Frameworks* (S. 401–408). Springer.

Stahnke, R., Schueler, S., & Roesken-Winter, B. (2016). Teachers' perception, interpretation, and decision-making: A systematic review of empirical mathematics education research. *ZDM*, *48*(1–2), 1–27.

Stein, M. K., & Lane, S. (1996). Instructional Tasks and the Development of Student Capacity to Think and Reason: An Analysis of the Relationship between Teaching and Learning in a Reform Mathematics Project. *Educational Research and Evaluation*, *2*(1), 50–80.

Verschaffel, L., Schukajlow, S., Star, J., & Van Dooren, W. (2020). Word problems in mathematics education: A survey. *ZDM*, *52*(1), 1–16.

GENDER-RELATED BELIEFS OF PROSPECTIVE MATHEMATICS TEACHERS

Georg Pfeiffer and Daniela Assmus

Martin Luther University Halle-Wittenberg

Gender-specific differences in mathematics to the disadvantage of girls are attributed, among other things, to stereotypical beliefs of teachers. However, it is not known to what extent such beliefs have already been developed in prospective mathematics teachers. In this article, a qualitative interview study with 23 german pre-service teachers is presented, which aimed to reconstruct gender-related beliefs. Beliefs about various gender-related differences as well as their causes could be reconstructed. Many participants showed an awareness of stereotypical role models and their causes for gender-related differences, but also own stereotypical beliefs. Conclusions, especially for mathematics didactics in teacher education, are drawn on the basis of empirical findings.

THEORETICAL BACKGROUND

A number of empirical studies have shown that in many countries, including Germany, gender differences in mathematical activities exist to the disadvantage of girls. Results of PISA 2000, IGLU-E 2001 or TIMSS 2007-2019 demonstrate a greater mathematical performance of boys than girls (Blunck & Pieper-Seier ,2010; Nonte et al., 2020). In addition to the poorer performance, studies show that girls encounter mathematics with a significantly lower interest in mathematics (Ganley & Lubienski, 2016), are less motivated to perform (Klieme et al., 2010) or show lower performance-related self-confidence in mathematical abilities than boys (Skaalvik & Skaalvik, 2004). Nowadays, such gender differences are mainly attributed to society and education (Blunck & Pieper-Seier, 2010). For example, societal influences include the stereotyping of mathematics as a male domain, whereas traditional gender role stereotypes are marked as unfavorable parenting styles in attitudes of parents (ibid.).

Likewise, the different expectations of teachers towards boys and girls influence gender-specific differences: Budde (2009), for example, demonstrated that teachers assume that boys in particular are mathematically gifted. Fischer and Rustemeyer (2007) argue that teachers expect girls to perform significantly less in mathematics, which leads to lower performance by girls. Butler (1994) also found that teachers who underestimate the abilities of female pupils are more likely to offer help to them and are more likely to show compassion towards them. In addition, studies show low awareness among teachers of the impact of their gender expectations on their students' achievement and ability self-concepts (Büker, & Rendtorff, 2015). Rather, the expectation-oriented teacher behavior is, among other things, an expression of gender-related (and partly implicit) beliefs that are stereotyped (ibid.).

2023. 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. 35-42). PME 46.

Beliefs are considered to be individual mental constructs that have an emotional component with an individual normative character (Hannula, 2012; Philipp, 2007) and develop in the course of engagement with the environment (Rokeach, 1968). In the case of teachers, acquired knowledge (e.g., during teacher education) and experiences (e.g., in teaching or in one's own past school years) are considered as central developmental factors for beliefs (Buehl & Fives, 2009). However, to our knowledge, nothing is known about the genesis of gender-related and possibly stereotypical beliefs.

Considerations to raise awareness of socially determined gender differences in the subject of mathematics already start in university teacher education. Suggestions for the conception of seminar series can be found, for example, in Langfeldt and Mischau (2011). The aim is to acquire gender competence as a "key competence related to the occupational field" (ibid., p. 315, translated by the authors). For example, prospective teachers should acquire abilities for the gender-sensitive design of teaching and learning ("didactic-methodical competence") and of interaction processes ("interactional competence"), scientific foundations with regard to gender research ("subject competence") and abilities to reflect on their own gender images and their own gender-related expectations ("Self competence"). However, for the training of prospective mathematics teachers it is criticized that the acquisition of gender competence is not anchored in Germany and therefore this acquisition of competence is of little or no importance at universities in Germany (ibid.).

RESEARCH INTEREST

The aim of the following study is to reconstruct gender-related beliefs of prospective mathematics teachers. The interest is based on the fact that while teachers' beliefs are considered to influence gender differences between boys and girls, little is known about the beliefs themselves (Büker, & Rendtorff, 2015). In this respect, it is of particular interest to what extent prospective mathematics teachers have already developed gender-related beliefs and whether these are stereotyped. Furthermore, the reconstruction of these beliefs is useful due to the possibility to derive and legitimize reflections on university teacher education with regard to a sensitization of gender-specific differences. University teacher training of prospective mathematics teachers in Germany hardly focuses on gender aspects (Langfeldt & Mischau, 2011). Therefore, it seems all the more important to explore gender-specific beliefs in order to break down stereotypical beliefs through targeted university educational programs, if necessary.

METHODOLOGICAL CONSIDERATIONS

A qualitative research and analysis method was chosen for the described research interest using a semi-standard interview.

Participants were 23 pre-service primary school teachers (female: n=19, male: n=4). These students were taking university courses in teacher education with mathematics

as a subject at a German university and were thus in the first phase of their professional education, in which hardly any practical teaching experience is gained in schools. Participants were at different stages of the teacher education program.

The conception of the interview is guided by basic principles of the problem-centered interview form according to Witzel and Reiter (2012). The aim of this interview form is to bring out the interviewees' subjective views on relevant problems by providing concrete conversation stimuli (ibid.). The problem-centered interview is therefore suitable for the reconstruction of beliefs (e.g., Eichler & Schmitz, 2018). In the present study the interviewees were asked to take a stand on whether boys and girls should be taught mathematics separately or why the proportion of girls in extracurricular activities to foster mathematical talent is low. In addition, various mathematical tasks of primary school lessons (e.g., factual, arithmetic and geometry tasks) were used as conversation stimuli for which the interviewees were asked to assess which of these tasks were particularly aimed at girls or boys. The tasks were chosen to address stereotypical beliefs regarding preferred topics or activities of the genders.

The evaluation of the data was carried out using qualitative content analysis according to Mayring (2014). After total transcriptions of the interviews passages which are substantial in content were found. Arguments and evaluations were regarded as substantial passages, as they refer to mental constructs of the interviewee that have an emotional component with an individual normative character and thus mark beliefs (Hannula, 2012; Philipp, 2007). This methodological approach on reconstructing beliefs is already established (Pfeiffer, 2023). From the reconstructed beliefs, a category system was then compiled based on Mayring (2014). The aim of this categorisation was to capture the complexity of the case-related analysis results or the different beliefs in order to make general statements about the types of gender-related beliefs of prospective mathematics teachers.

RESULTS

A category system of gender-related beliefs

From the data, beliefs about the following deductively and inductively developed categories could be reconstructed:

1. gender disparities....

    a....with regard to mathematics achievement (e.g., "It is usually the case that boys are gifted in maths and girls flourish in languages.")

    b....in terms of subject-related interests (e.g., "Boys are more interested in mathematics than girls.")

    c....in terms of preferred mathematical activities and tasks (e.g., "Girls are attracted by tasks where they can draw/colour something.", "For boys, things that have a clear structure with a fixed sequence are worthy of consideration. Girls tend to be puzzlers and then they deal more intensively with another direction.")

d....with regard to affective characteristics (e.g., “Girls do not dare to participate in extracurricular maths projects because they are afraid of being laughed at if they ask a question.”)

e....with regard to behavioural characteristics (e.g., “Boys are bold and sometimes give the wrong approach, while girls think longer and think through backgrounds.”)

2. explanatory approaches to gender-related disparities via socialisation-related factors / stereotypical ideas ...

a....among parents (e.g., “There is a popular idea among parents that boys are more gifted in science subjects than girls.”)

b....among teachers (e.g., “Gender differences in mathematics lessons are caused by teachers because girls feel put off if only boys are praised.”)

c....in society (e.g., “It is socially influenced to say that boys are better at mathematics and science than girls”, “Girls think they are worse in mathematics because this prejudice is transmitted through generations.”)

3. non-existence of gender-specific disparities ...

a....with regard to mathematics achievement (e.g., “Girls can do maths just as well as boys.”)

b....in terms of preferred mathematical activities (e.g., “Boys do not prefer different mathematical activities than girls because it depends on how interested you are in the subject and not whether you are a boy or a girl.”)

Analysis and discussion of selected aspects

Overall, the pre-service teachers’ statements and the reconstructed beliefs largely fit well with study results on mathematics-specific gender-related gaps between boys and girls and their causes, so that many categories resulted deductively from theoretical foundations. If mathematics-specific disparities were mentioned, they mostly related to interest in the subject. Thus, the majority of the participants attested boys a greater interest in the subject of mathematics than girls. Differences in performance were only rarely mentioned and then usually justified by a low interest in the subject or equated with it. With regard to preferred mathematical activities and tasks, most reconstructed beliefs did not assume any general gender-related differences. Nevertheless, stereotypical attributions were repeatedly expressed, such as drawing tasks in geometry or inventing arithmetic stories as preferred tasks for girls or purely symbolic arithmetic tasks for boys. Typically, however, students independently labeled these as stereotypical views or prejudices.

Socialisation-related factors were predominantly mentioned as causes for possible differences between girls and boys, which in turn were attributed to stereotypical ideas in society in many cases. Thus, many students expressed an awareness of the importance of stereotypical ideas. Stereotypical ideas were mainly attributed to parents

or to society in general. Teachers were rarely addressed in this context. One example is represented by Ms. B., who responded to the question of how to increase the proportion of girls in extracurricular projects for mathematically interested and gifted children:

> "Yes (.) so first of all avoid these stereotypical phrases then (.) out of that also if now for example if I now as a teacher stand in front of the class really address all the children and not just yes you boys you are always so good or so ((laughs)) but yes address everyone."

The infrequent focus on teachers may indicate that the interviewees were previously unaware of influences of stereotypical beliefs among teachers on gender differences. This empirical finding is also reflected in other studies (Büker et al., 2015). On the one hand, this result can be explained by a still missing identification of the students with their future teaching profession. They may also have been unaware of their own role in causing gender differences due to a lack of teaching experience. Possibly, corresponding beliefs first develop with experience in everyday school life (Buehl & Fives, 2009). Nevertheless, increased awareness of gender inequality in teacher education can contribute to pre-service teachers' knowledge of their own role, which in turn can shape the genesis of beliefs in the future teaching profession. On the other hand, the result can be attributed to the fact that the societal or socialization-related stereotyping of mathematics as a male domain (Blunck et al., 2010) is recognized but also accepted by the students. Possibly, through unilateral explanations of gender differences, pre-service teachers evade their later pedagogical duty to counteract inequality with regard to the category of gender. Thus, the development of comprehensive gender competence in teacher education is relevant.

Even though the beliefs of many pre-service teachers point to an awareness of the stereotyping of mathematics as a male domain, various interview passages can be identified in which stereotypical beliefs can be suspected. These are found in interview passages that do not explicitly ask about differences between genders. Two examples will illustrate this.

When asked about differences in terms of preferred mathematical activities, Ms. H. does not want to make blanket attributions:

> "I think that's pretty individual, so I don't think you can say that across the board. I don't know if there's something that boys generally like better than girls, I think it depends on the person."

At the end of the interview, Ms. H. is asked whether she considers gender-segregated teaching to be useful in the subject of mathematics. She rejects this. She answers the question of what this could look like if it were to be used as follows:

> "Maybe with the girls it would be even more with painting somehow I can imagine now so just something geometric and with the boys maybe to what extent that's possible something crafty maybe even, something you can link with math, something like that."

It is not possible to assess whether Ms. H. is aiming at suspected differences in performance or interest.

Mr. H. has met many girls with good math grades in his own school biography. In his opinion, girls and boys do not differ in terms of mathematics performance. He rejects separate mathematics classes:

> "No. I have to say that, because everyone benefits from learning together [...] Girls could also take over an enthusiasm from the boys, so to speak "look how fast he can do that, I want that too" and something like that. I would totally reject that, because all children have the right to learn together somehow, I think, and also to learn from each other, so I wouldn't advocate teaching that separately."

The focus on the boys' speed indicates that differences are implicitly assumed here as well.

Similar statements can be found in further interviews. They suggest that in situations of quick decisions, the often existing awareness of stereotypical ideas and their consequences does not take effect, but that the nevertheless existing unconscious stereotypical ideas guide actions. In order to cope with the challenges in everyday pedagogical life, the category of gender is used. Gender-related and stereotypical beliefs serve here to filter complexity. It can be assumed that this does not reduce gender-related differences, but rather reinforces them.

SUMMARY AND CONCLUSIONS

The study provides empirical findings on gender-specific beliefs of pre-service teachers. With the help of formed categories it was possible to bundle the aspect variety of the beliefs. The central findings are that the gender-related beliefs (1) point to a greater interest in mathematics among boys, (2) point to an awareness of the importance of stereotypical views among parents and society in general, (3) hardly refer to the role of teachers as contributors to gender differences, and (4) are partly stereotype-loaded. The infrequent focus on teachers and the partly stereotype-loaded beliefs legitimize to increasingly address gender as a dimension of discrimination in teacher education.

As a consequence of the results, we see the function of gender-sensitive courses primarily in the support of pre-service teachers in the development of gender-sensitive action competence. Following the conceptual considerations of Langfeld and Mischau (2011) for the design of gender-sensitive courses using the competence dimensions "subject competence", "didactic-methodical competence", "interactional competence" and "self-competence" to be developed, we locate suitable activities for the development of a competence to act above all in the examination of didactic-methodical questions and in the examination of interactions in mathematics teaching. For the latter, a differentiated analysis of conversation excerpts from classroom situations with regard to stereotypical interaction patterns is the first step. An essential element is the joint development of alternative, gender-neutral language patterns and reactions. Role plays, for example, can be used to incorporate these into the participants' own repertoire of actions. It would make sense to deepen these

experiences in the school class and also to analyse and reflect on one's own interaction patterns.

In addition, other research interests emerge from the results of the study. It can be seen that prospective mathematics teachers have gender-specific and, in some cases, stereotypical beliefs. However, nothing is known about the development of such beliefs. It is therefore interesting to find out, first, whether there is a connection between views of mathematics or of mathematics teaching and gender-specific beliefs. Grigutsch et al. (1998), for example, distinguish formalistic, schematic, process-oriented, and application-oriented views of mathematics and mathematics teaching. Second, it might be interesting to investigate a connection between educational biography in mathematics and gender-specific beliefs. Possibly, the beliefs are shaped by their own learning success in mathematics.

REFERENCES

Blunck, A., & Pieper-Seier, I. (2010). Mathematik: Genderforschung auf schwierigem Terrain. In R. Becker & B. Kortendieck (Eds.), Handbuch Frauen- und Geschlechtserforschung. Theorie, Methoden, Empirie (pp 820-828). VS-Verlag.

Buehl, M. M., & Fives, H. (2009). Exploring teachers' beliefs about teaching knowledge: Where does it come from? Does it change? Journal of Experimental Education (77), pp. 367–407.

Büker, P., & Rendtorff, B. (2015). Sichtweisen von Lehrkräften auf Ethnizität und Geschlecht. Eine Problemanzeige. Zeitschrift für Pädagogik 61, (1), pp. 101-117.

Budde, J. (2009). Mathematikunterricht und Geschlecht. Empirische Ergebnisse und pädagogische Ansätze. In: Bundesministerium für Bildung und Forschung (BMBF) (Eds), Band 30.

Butler, R. (1994). Teacher communications and student interpretations: Effects of teacher responses to failing students on attributional inferences in two age groups. British Journal of Educational Psychology, 64, pp. 277–7294.

Eichler, A., & Schmitz, A. (2018). Domain specificity of mathematics teachers' beliefs and goals. In B. Rott, G. Törner, J. Peters-Dasdemin, A. Möller Safrudiannur, Views and beliefs in mathematics education. The role of beliefs in the classrooom. (pp. 137-146), Springer.

Fischer, N., & Rustemeyer, R. (2007). Motivationsentwicklung und schülerperzipiertes lehrkraftverhalten im mathematikunterricht. *Zeitschrift für Pädagogische Psychologie / German Journal of Educational Psychology, 21*(2), pp. 135-144.

Ganley, C. M., & Lubienski, S. T. (2016). Mathematics confidence, interest, and performance: Examining gender patterns and reciprocal relations. Learning and individual differences, 47, pp. 182-193.

Grigutsch, S., Raatz, U., & Törner, G. (1998). Einstellungen gegenüber Mathematik bei Mathematiklehrern. Journal für Mathematik-Didaktik 19 (1), pp. 3–45.

Hannula, M. S. (2012). Exploring new dimensions of mathematics-related affect: embodied and social theories. Research in Mathematics Education 14 (2), pp. 137–161.

Klieme, E., Schümer, G., & Knoll, S. (2001). Mathematikunterricht in der Sekundarstufe I: „Aufgabenkultur“ und Unterrichtsgestaltung im internationalen Vergleich. In E. Klieme, & J. Baumert (Hrsg.), TIMSS – Impulse für Schule und Unterricht (pp. 43–57). BMBF.

Langfeldt, B., & Mischau, A. (2011). Genderkompetenz als Bestandteil der Lehramtsausbildung im Fach Mathematik – zu innovativ für deutsche Hochschulen? ZFHE, 6(3), pp. 311-324.

Mayring, P. (2014). Qualitative content analysis: theoretical foundation, basic procedures and software solution. http://nbn-resolving.de/urn:nbn:de:0168-ssoar-395173.

Pfeiffer, G. (2023). Beliefs von Lehrkräften zum Einsatz von Arbeitsmitteln im arithmetischen Anfangsunterricht für Kinder mit Förderbedarf im Bereich Lernen. WTM.

Philipp, R.A. (2007). Mathematics teachers’ beliefs and affect. In F. K. Lester (Ed.), Second handbook of research on mathematics teaching and learning (pp. 257-315). Information Age Publishing.

Rokeach, M. (1968). Beliefs, attitudes, and values. A theory of organization and change. Jossey-Bass.

Skaalvik, S., & Skaalvik, E.M. (2004). Gender differences in math and verbal self-concept, performance expectations, and motivation. Sex Roles 50 (3/4), pp. 241-252.

Vester, M. (2013). Das schulische Bildungssystem unter Druck: Sortierung nach Herkunft oder milieugerechte Pädagogik ? In F. Dietrich & M. Thieme (Eds.), Bildungsgerechtigkeit jenseits von Chancengleichheit. Theoretische und empirische Ergänzungen und Alternativen zu ‘PISA’. Wiesbaden: Springer.

Witzel, A., & Reiter, H. (2012). The problem-centred interview: principles & practice. Sage.

REPRESENTING COVARIATION FUNCTIONAL SITUATIONS IN A TABLET-ENABLED DIGITAL LEARNING ENVIROMENT

Marios Pittalis[1], Eleni Demosthenous[1], and Ute Sproesser[2]

[1]University of Cyprus, [2]University of Ludwigsburg

The purpose of this study was to empirically investigate whether a tablet-enabled digital learning environment facilitates students to graphically represent functional situations that require the coordination of two co-varying quantities. Fifty-nine 11-year-old-students were asked to construct graphs for four scenarios. The results show that a significant percentage of students utilized the feedback provided by the learning environment to revisit their initial graph. The feedback involved watching how the animation matched the constructed graph, which helped students in overcoming typical mistakes, such as sketching lines that showed variation in only one of the quantities. Several students provided comprehensive explanations to match the graph with the scenario that reflected different categories of covariational reasoning.

INTRODUCTION AND THEORETICAL BACKGROUND

Researchers still struggle to find effective ways of integrating technology into mathematics teaching and learning. This also holds for the domain of functional thinking and graphical reasoning. It is important to know how portable and handheld digital technologies offer opportunities for enacting embodied learning in situations that require representing functional relations (Abrahamson & Bakker, 2016). This is also supported by the National Council of Teachers of Mathematics (NCTM, 2000) that emphasized both the use and connection of representations in making sense of functional relationships. The present study focuses on the aspect of connecting a physical situation with a graphical one by using digital tools / technologies. The goal of this study is to empirically investigate whether a tablet-enabled digital learning environment facilitates students representing graphically functional situations. We explored 11-year-old students representing distance changing over time in a variety of physical situations that require coordination of the co-varying quantities while using a table-enabled digital environment.

Functional Thinking

Functional thinking has been described as a key component in mathematics education and has generally been defined as the process of building, describing, and reasoning with and about functions. The covariation approach to functional thinking has recently received increasing attention in mathematics education and has been examined through contextualized dynamic functional situations (Thomson & Carlson, 2017). It has been linked to a dynamic view of the function concept as it supports a view of mathematics as a way of making sense of the phenomena of relationships of dependence, causation, interaction, and correlation between varying quantities. In this study, we conceptualize

2023. 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. 43-50). PME 46.

primary school students' functional thinking as the type of thinking that focuses on the invariant relation between two varying quantities and involves noticing, generalizing, and abstracting relations between covarying quantities; and representing these relations (Pittalis et al., 2020). Hence, representing the graph of a dynamic functional relationship is considered an important functional thinking ability. It requires reasoning about relatives changes of the involved quantities, the direction of these changes, and the translation between natural language and graphical representations to model dynamic functional situations.

Covariation and Graphical Reasoning

Students' ability to reason with quantities and relationships has empirically proved to foster functional thinking (Ellis, 2011). As proposed by Thompson and Carlson (2017), the covariation approach to function is supported by the covariational reasoning framework. This framework suggests that students should be able to think about two varying quantities and understand that this variation takes place simultaneously. This type of reasoning is important for interpreting and constructing graphical representations because it facilitates making connections between the involved quantities. The term of graphical reasoning encompasses both graph interpretation and construction. More precisely, graph interpretation includes visualizing features of a graph, interpreting relationships and connecting the identified relationships with what the graph represents (Shah & Hoeffner, 2002). For distance-time situations, students should be given the opportunity to connect the represented physical situation with visual elements of the graphical representation and vice versa. To do so, two instructional approaches have been suggested (Duijzer et al., 2019). The first one emphasizes the quantitative or local aspects of graphing, while the second one highlights the importance of grasping qualitative or global aspects of the graph. The global perspective is important in developing understanding of the real-life situation that is represented by the general shape of the graph. It facilitates visualizing the global relationship between two quantities before constructing an exact graph. Research shows that students encounter difficulties when representing changes over time, such as interpreting a graph as an iconic representation of a real situation or plotting and connecting points in the graph without considering the values between successive points and consequently without understanding that the graph represents a relationship between covarying quantities (Moore et al., 2013).

Embodied Learning Environments for Graphing Motion

Digital environments for graphing motion can be categorized in respect to bodily involvement and immediacy. Bodily involvement can be distinguished between own motion and observing others/objects' motion. Immediacy is defined in terms of immediate and non-immediate. An immediate task provides a simultaneous interaction with the physical environment, whereas in the second case this interaction is based on an embodied simulation. Embodied learning environments are effective when students make use of their own motion and link this motion immediately to a graphical

representation (Duijzer, et al. 2019). In this study, we explored students' understanding of graphing motion as they observed and influenced an objects' motion in which the bodily involvement took place in the absence of direct stimuli.

THE PRESENT STUDY

The purpose of the present study was to investigate students representing qualitative and quantitative aspects of co-varying quantities (distance-time scenarios) in a tablet-enabled learning environment. In line with the lenses of the above theoretical frameworks, the research questions are: (a) Does a tablet-enabled learning environment that provides students an interactive graphing motion tool facilitate 11-year-old students to represent qualitative and quantitative aspects of distance-time scenarios and (b) do students use covariational reasoning while representing distance-time scenarios?

Participants, Intervention, Tasks and Procedure

Fifty-nine 11-year-old-students (28 girls and 31 boys) from three Grade 5 classes participated in the study. They had used tablets in their mathematics lessons several times before. The students had not been taught functional situations involving measures that covariate simultaneously. In line with the national curriculum guidelines, students mainly focus on exploring patterns and function tables.

The intervention was designed in the framework of the Erasmus+ project FunThink with the goal to foster functional thinking. For this purpose, a module of four 40-minutes lessons was designed with an emphasis on the aspect of covariation. The lessons were delivered by a member of the research team in a two-week period. The module includes activities that require conceiving co-varying quantities, representing graphically distance-time scenarios, animating physical movement scenarios and distance-time graphs with an applet. All students had access to the applet on tablets. Based on the covariation aspect of function, we used the online applet *Turtle Crossing* from the Desmos platform (desmos.com) in the intervention. The applet presents a turtle that walks away from the sea and students make connections between several turtle-crossing scenarios and graphs. We used the functionality to draw a distance-time graph and then watch an animation of the turtle's journey (see Table 1).

Two weeks after the completion of the intervention program, each student was interviewed in a session of approximately 20 minutes. We provided four scenarios and students had to construct a graph in the *Turtle Crossing* applet that corresponds to the given scenario. Students were informed that they could edit their work based on the feedback provided by the animation that showed the turtle's journey according to the student's graphs. A warm-up example familiarized the students with the procedure. The researcher prompted the students to validate their work by getting feedback from the animation and raised questions to investigate if students could match parts of the graph with specific moving actions. Two scenarios required conceptualizing global aspects of the graph while the other two involved grasping local aspects (see Table 1).

Scenarios - Global	Scenarios - Local

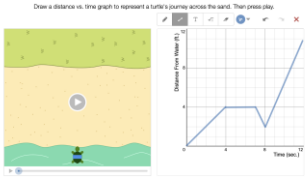
Task 1: The turtle moves away from the sea. Suddenly, it stops for a while. Then, it continues moving away from the sea. **Task 2:** The turtle moves away from the sea. Suddenly, it stops for a while. It then starts returning to the water. Before reaching the water, it decides to move away again.	**Task 3:** The turtle moves 8 ft away from the sea in 4 seconds. It pauses for 2 seconds. It then returns to the sea in just 2 seconds. **Task 4:** The turtle is 4 ft away from the sea. It moves 2 ft away over the next 4 seconds. It then returns to the water by traveling 2 ft per second.
 desmos.com	

Table 1: Description of Tasks

Data collection and Analysis

This study used qualitative methods for data collection and analysis. We videotaped the interviews to capture students' work on the tablets as well as their verbal explanations. A qualitative interpretive framework was used for the data analysis (Miles & Huberman, 1994), which consisted of three phases. The first phase focused on examining whether the constructed graph matched the scenario, and whether the student could explain which part of the graph corresponds to scenarios. In the second phase, we examined students' mistakes and the extent to which they utilized the feedback provided by the applet to edit their work. Finally, we analyzed students' explanations to identify covariation reasoning instances. To do so, we set-up a coding framework that synthesized theory and data driven codes.

RESULTS

We addressed the first research question of the study by examining students' graph constructions and their verbal explanations. Particularly, we analyzed whether students' graphs corresponded to the given scenario and whether they revised appropriately their graphs based on the feedback provided by the applet.

Table 2 presents the number of students that constructed the graph, which corresponded to the given scenario from their first attempt and the number of students that used the feedback from the applet to revise their graph. In Task 1, 46% of the students constructed the correct graph from their first attempt, while in Task 2 the respective percentage was 54%. Both tasks involved understanding of co-varying functional situations at a global level. In Task 1, 46% of the students effectively used the provided feedback by the animation and revised their constructed graph after one/two or multiple attempts. To inspect incorrect solutions in more detail, typical student mistakes were categorized. These mistakes appeared in their first attempt or after multiple attempts. Several students represented the turtle's pause by leaving a blank space between two segments (see Table 3, Task 1(a)) or by missing to represent it (Table 3, Task 1(c)).

Other students constructed a vertical segment at some part of the graph, indicating difficulty in handling the independent variation of time (Table 3, Task 1(b) and (d)). In Task 2, 37% of the students managed to revise their graph based on the provided feedback. In Table 3, Tasks 2(a)-(c) present students' difficulty in constructing the turtle's return to the water by handling both the increase of seconds (time) and the decrease in feet (distance). Also, several students faced difficulties in handling precisely the quantity that shows the turtle's distance from the water by representing that the turtle changed direction before reaching the sea (Task 2(d)).

	Correct from the first attempt	Correct after one/two revisions	Correct after multiple revisions	Incorrect after multiple attempts
Task 1	27/59 (46%)	12/59 (20%)	15/59 (26%)	5/59 (8%)
Task 2	32/59 (54%)	15/59 (26%)	7/59 (11%)	5/59 (9%)
Task 3	22/59 (37%)	12/59 (20%)	17/59 (29%)	8/59 (14%)
Task 4	2/59 (3%)	17/59 (29%)	8/59 (14%)	32/59 (54%)

Table 2: Correctness of students' graphs across tasks

In the local scenarios, students' graph constructions were less successful compared to the global ones. In Task 3, 37% of the students and in Task 4, only 3% managed to create a graph that corresponded to the given scenario from their first attempt. However, 49% of the students revised their graph correctly in Task 3 based on the animation after one/two or multiple attempts. As shown in Table 3, Tasks 3(a)-(c), students faced difficulties in representing the turtle's returns to the sea (the distance to the water decreases) while the time is still running (the time in seconds increases). It seems that the provided feedback helped students to identify their mistake and revise it accordingly. Further, students faced difficulties in representing the time needed for the turtle to return to the water (Task 3(d)). In Task 4, 43% of the students managed to revise their graph successfully. It was common among students to interpret the phrase "the turtle is 4 ft away from the sea" as the turtle moves 4 ft away from the sea in 4 seconds (see Task 4(a) in Table 3). Hence, they tended to use the point (0,0) as the starting point instead of the point (0,4). Other students sketched a vertical segment on the y-axis from the origin to the point (0, 4), while others sketched a horizontal segment from the point (0, 4) to the point (4, 4) (see Tasks 4(b) and 4(c)). Again, some students represented the scenario "It moves another 2 ft away over the next 4 seconds" by using a vertical segment, ignoring the increase of time. The students also faced difficulties in representing the last part of the scenario, since it provided the rate of distance-time. Twelve out of the 19 students that succeeded in the task, followed a step-by-step strategy, constructing the return second by second till the turtle reached the x-axis.

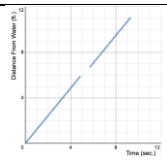
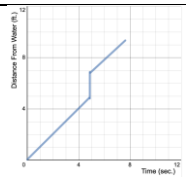
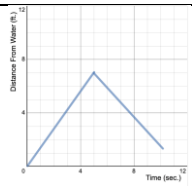
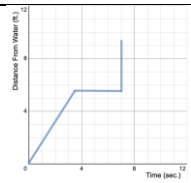
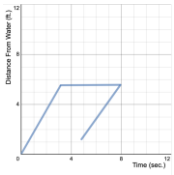
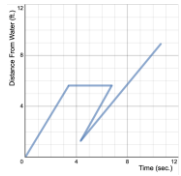
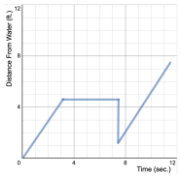
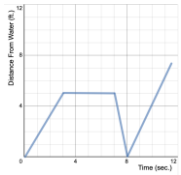
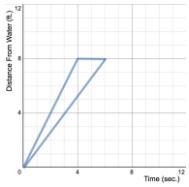
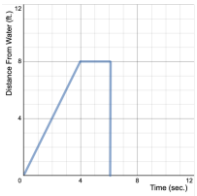
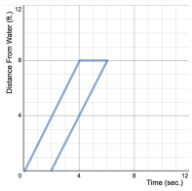
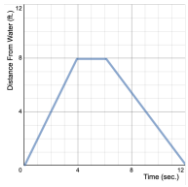
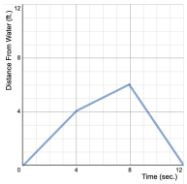
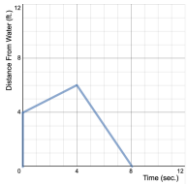
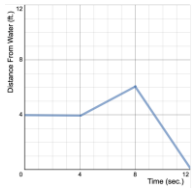
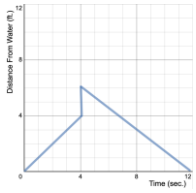
	(a)	(b)	(c)	(d)
Task 1				
Task 2				
Task 3				
Task 4				

Table 3: Typical mistakes of students in the four tasks

To answer the second research question, we examined students' explanations while working on the four tasks. We analysed how they explained the correspondence between the constructed graph and the scenario, and how they reasoned about the time intervals for specific parts of the scenarios in Tasks 3 and 4 (see Table 4). This analysis provided evidence regarding students' covariation reasoning. In Tasks 1 and 2, 60% and 65% of the students respectively, provided correct interpretation of the graph by explaining the correspondence of each part of the graph with the given scenario. Although it was not directly required, in Task 1, 33% of the students that provided a part-part explanation and in Task 2, 20% included numeric data (i.e., concrete values of the quantities) in their explanations. Further, 21% and 26% of the students provided general or partially correct explanation based on the animation or the graph, without making clear connection with the scenario in Tasks 1 and 2, respectively. In Task 3, 28% of the students provided general or partially correct explanation based on the animation or the graph, while 58% explained adequately the correspondence between the graph and the scenario. It is worth mentioning that half of the students that provided part-part explanation relied on the aspect of time when comparing the velocity of the turtle in the parts of the journey, while the other half coordinated time and distance. In Task 4, only 12% of the students provided general or partially correct explanation based on the animation or the graph and 43% provided an adequate explanation that can be categorized into 3 groups. The first group (5 out of 25) explained the format of

their graph without relating explicitly the two quantities. The second group (12 out of 25) grasped the notion of rate for the return of the turtle, without relating the rate of 2 feet per second with the total distance that had to be covered. The third group (8 out of 25) coordinated effectively the two-covarying measures, as they calculated in advance the time needed by the turtle to return based on the given rate and calculated the rate of feet per second for different parts of the graph.

Therefore, we identified six categories of covariational reasoning in students' representations of covariation situations, based on the explanations they provided in the four tasks (Table 4): (a) Lack of coordination of two quantities graphically, (b) qualitative coordination in one (or selected) dimension (e.g. interpretation of moving away from the water = only as one quantity increases, the other quantity also increases but not being able to sketch that the turtle stays still), (c) qualitative coordination in multiple dimensions, (d) quantitative coordination in one dimension (e.g. interpretation of walking 8ft in 4 s = as one quantity increases by 8ft, the other quantity increases by 4s), (e) quantitative coordination in multiple dimensions, and (f) abstract coordination (can compare by unitization, can calculate time/distance needed for given rate).

	Part-part explanation	General or partial explanation based on the animation	General or partial explanation based on the graph	No explanation or incorrect interpretation
Task 1	35/59 (60%)	7/59 (11%)	12/59 (20%)	5/59 (9%)
Task 2	39/59 (65%)	5/59 (9%)	10/59 (17%)	5/59 (9%)
Task 3	34/59 (58%)	8/59 (14%)	8/59 (14%)	9/59 (14%)
Task 4	25/59 (43%)	2/59 (3%)	5/59 (9%)	27/59 (45%)

Table 4: Students' explanations across tasks

DISCUSSION

The contribution of this study lies on the empirical examination of the way 11-year-old students graphically represent functional situations that require coordination of two co-varying quantities in a tablet-enabled digital learning environment that provides an animation of the constructed graph. The results of the study show that in the two tasks with the global aspects, around half of the students constructed a correct graph in their first attempt and about 40% of the students correctly edited their work and provided appropriate explanations that matched each part of the graph with the scenario, by using the feedback provided by the learning environment in the form of an animation of the turtles' journey. The provided feedback facilitated overcoming typical mistakes in representing covarying quantities such as coordinating the variation of the involved measures and interpreting a graph as an iconic representation (Moore et al., 2013). The most typical one was representing the increase of time as the distance stays stable or decreases (Shah & Hoeffner, 2002). The percentage of students that constructed a correct graph that involved representing local aspects in the first attempt was smaller,

but again, about half of the students managed to edit appropriately their graph based on the provided feedback. To do so, several students coordinated quantitatively only one of the involved quantities, others coordinated quantitatively both quantities and some of them exhibited an abstract coordination of the co-varying quantities by unitizing and calculating time/distance for different parts of the graph. The results of the study empirically show the potential of integrating tablet-enabled learning environments in mathematics teaching to develop the covariation aspect of functional thinking and graphical reasoning. Further, the study shows possible ways to utilize the embodied nature of distance-time activities to model dynamic functional situations by making self-dynamic connections between scenarios and graphical representations.

ACKNOWLEDGMENT

This study is co-funded by the Erasmus+ program of the European Union. The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

REFERENCES

Abrahamson, D., & Bakker, A. (2016). Making sense of movement in embodied design for mathematics learning. *Cognitive research: principles and implications, 1*(1), 1-13.

Duijzer, C., den Heuvel-Panhuizen, V., Veldhuis, M., Doorman, M., & Leseman, P. (2019). Embodied learning environments for graphing motion: A systematic literature review. *Educational Psychology Review, 31*(3), 597-629.

Ellis, A. B. (2011). Algebra in the middle school: Developing functional relationships through quantitative reasoning. In J. Cai & E. Knuth (Eds.), *Early algebraization: A global dialogue from multiple perspectives* (pp. 215–238). Springer-Verlag.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis* (2nd ed.). Sage.

Moore, K. C., Paoletti, T., & Musgrave, S. (2013). Covariational reasoning and invariance among coordinate systems. *The Journal of Mathematical Behavior, 32*(3), 461-473.

National Council for Teachers of Mathematics. (2000). *Principles and standards for school mathematics.* Author.

Pittalis, M., Pitta-Pantazi, D., & Christou, C. (2020). Young students' functional thinking modes: The relation between recursive patterning, covariational thinking, and correspondence relations. *Journal for Research in Mathematics Education, 51*(5), 631–674.

Shah, P., & Hoeffner, J. (2002). Review of graph comprehension research: Implications for instruction. *Educational Psychology Review, 14*(1), 47-69.

Thompson, P. W., & Carlson, M. P. (2017). Variation, covariation, and functions: Foundational ways of thinking mathematically. In J. Cai (Ed.), *Compendium for research in mathematics education* (pp. 421–456). National Council of Teachers of Mathematics.

A META-DISCIPLINARY REFLECTION ON A STEAM SCHOOL ACTIVITY: THE ROLE OF MATHEMATICS

Gabriella Pocalana, Ornella Robutti and Giulia Bini

Università degli Studi di Torino

This paper aims at providing a meta-disciplinary reflection on the role of mathematics in the design of a STEAM activity for lower-secondary schools. The analyzed case study involves a group of teachers of different disciplines from the same Italian school, who collaboratively designed an interdisciplinary STEAM activity, following a template created by researchers in mathematics education, in the context of an international European project. We analyze data gathered in an interview with the mathematics teacher of the group, conducted to gain insight into the role of mathematics in the design of the activity, from her point of view. Different roles emerged from the teacher's reflections about the subsequent phases of the project, entailing an evolving relationship between mathematics and the other disciplines.

INTRODUCTION AND THEORETICAL BACKGROUND

Interdisciplinarity indicates an approach to a topic that embraces the competences of different scientific sectors or disciplines, which cooperate to provide a wider understanding of a common topic (Capone, 2022). More than 50 years ago, at the Seminar on Interdisciplinarity in Universities held in Nice (France), scholars joined efforts to promote interdisciplinarity and identify the role of mathematics in interdisciplinary activities, reckoning that interdisciplinarity "is important to allow students to find themselves in the present-day world, to understand and criticize the flood of information they are deluged with daily" (Apostel et al., 1972, p. 14). This becomes even more relevant today, in a world that increasingly demands intertwined skills to make sense of its complexity. Indeed, interdisciplinarity in mathematics education (Doig et al., 2019) and in STEAM (Science, Technology, Engineering, Arts and Mathematics) education (Khine & Areepattamannil, 2019; Henriksen, 2019) has been assuming increasing importance in the last decades. Yet, Piaget warns us that, due to its deductive nature, mathematics has a "particular independence" (Piaget, 1972, p. 127) from other scientific disciplines of experimental nature. This peculiarity requires a focused effort by the teachers to make interdisciplinary connections between mathematics and other disciplines apparent to students.

Researchers in mathematics education in many parts of the world implement "various sorts of conjunction of mathematics with other knowledge" (Williams & Roth, 2019). In these implementations, the role of mathematics is interpreted in different ways (Stohlmann, 2018). Sometimes, mathematics is conceived as a tool for doing the computational work needed to solve scientific/technological problems; in other cases, mathematics is used to model phenomena observed in other disciplines (Williams & Wake, 2007).

2023. 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. 51-58). PME 46.

The aim of this paper is to conduct a *meta-disciplinary* (Williams & Roth, 2019) reflection on the role of mathematics in the design of an interdisciplinary STEAM activity for lower-secondary schools, based on how it is perceived by the mathematics teacher as a co-designer of the activity. Addressing this issue, we can contribute to deepening the insight into interdisciplinary STEAM teaching and collaborative design of STEAM activities by teachers of different disciplines. Our research question is: What is the role of mathematics in a STEAM activity, cooperatively designed by teachers of different disciplines, according to the mathematics teachers' perspective?

CONTEXT AND METHODOLOGY

Our study context is the Erasmus+ Project named STEAM-Connect, where researchers from different countries (Austria, Finland, Italy, Luxembourg and Slovakia) and teachers of different disciplines (arts, mathematics, music, technology, etc.) work collaboratively to design STEAM activities for all school grades. Different levels of connection are expected as outcomes of the project: a connection between teachers of different disciplines in each school, between teachers of different schools in the same country; between researchers and teachers of the same country, between researchers from different countries, and between researchers and teachers from different countries during the dissemination phase. We claim that the process of designing an interdisciplinary STEAM activity is far from trivial. It takes a long time and multiple design cycles for teachers to learn how to collaborate and coordinate to choose a common object (a problem or a topic) and to pursue the common aim of the activity.

The STEAM-Connect project has a duration of three years and started in November 2021. During the first months of the project, researchers from all the involved countries co-created a common template for the STEAM activities to be designed by teachers. The common template had the aim of supporting the exchange, sharing and dissemination of the STEAM activities within different schools, in the same country or in different countries, in the final phase of the project.

The data analyzed in this paper refer to the Italian national phase of the project, in which the teachers used the template provided by the researchers to collaboratively design STEAM activities. For our investigation, we selected as a case study a lower-secondary school in Piedmont, open to educational innovation and already engaged with the Mathematics Department of the University of Turin (UNITO) in other projects. Firstly, we analyzed the template filled in by the teachers of this school to gain an overview of the STEAM activity in terms of involved disciplines, school grade, time needed and learning objectives. Then, we conducted a semi-structured interview with the mathematics teacher of the school, Paola (pseudonym), who is a very experienced teacher, having also the role of teacher-educator in other projects. She answered two main questions, which constituted the baseline of the interview: 1) "How did the idea of the STEAM activity that you designed in your school come about?"; 2) "How would you describe the role of mathematics in the STEAM activity that you designed in your school?"

With a qualitative methodology and an interpretative approach (Cohen et al., 2007), we analyzed the transcript of Paola's interview, to answer our research question, with specific attention to the evolution of the role of mathematics in the design of the STEAM activity during the project, as described by Paola.

RESULTS

From the first part of the template designed by the researchers (Figure 1, translated from Italian by the authors), filled in by Paola and her colleagues to describe their STEAM activity, we understand that it is intended for a sixth-grade class and it deals with the topic of symmetries and translations. The involved disciplines are: arts, music, technology, science and mathematics. Among the learning objectives, the teachers declare a general objective, common to art, music and science, which refers to the real-world applications of the concept of symmetries and translations. Besides that, they declare a learning objective specific to mathematics, connected with the mathematics curriculum: the representation of symmetries and translations in the Cartesian plane.

In the interview (translated from Italian by the authors), Paola explains the path which led her and her colleagues to the design of the STEAM activity described in the template. She reports that the idea for the original nucleus of the activity, emerging at the very beginning of the STEAM-Connect project, involved only two teachers and three disciplines: arts, science and mathematics (these last two taught by Paola herself).

Paola proposes an interesting reflection on the role of mathematics as the main aim of the activity: arts and science were not involved *per se*, but as opportunities to talk about symmetries, starting from concrete examples. Indeed, the activity was carried out during the mathematics class, with the art teacher assuming the role of "special guest".

> Paola: The first collaboration with the art teacher […] entailed starting from the study of symmetries in artworks and in natural elements, to arrive at the discovery of the Fibonacci sequence and the golden rectangle. […] in that case, it was mathematics, let's say, that dominated science and art and it was the art teacher who came in co-presence with me to carry on this activity.

DISCIPLINES	GRADE	TOTAL ACTIVITY TIME	LEARNING OBJECTIVES DURING THE LESSON - DISCIPLINARY COMPETENCIES	LEARNING OBJECTIVES AFTER THE LESSON - NON DISCIPLINARY COMPETENCIES
• Mathematics • Science • Arts • Music • Technology	6th grade	• 12 hours (at school) • 3 hours (at home)	• Building *moduli* with different materials. • Applying the axial symmetry in a drawing. • Representing symmetrical figures in the Cartesian plane. • Identifying symmetries in the piano keyboard. • Recognizing a translation in a polyphonic piece in canon form. • Performing a canon song.	• Recognizing symmetries and translations in art works, music and natural elements (landscapes, flowers, fruits, plants …). • Interpreting images, sounds and movements in a "modular" key • Representing symmetries and translations on the Cartesian plane.

Figure 1. First part of the STEAM activity template.

A similar dynamic is reported by Paola regarding the relationship between mathematics and music, during the first collaboration of the teachers of these disciplines on the topic of symmetries. She explains that the idea of the topic came from a teacher professional development program (SSPM), which she followed at the Mathematics Department of the University of Turin (Pocalana & Robutti, 2022).

Paola: During the SSPM professional development program, they proposed an activity on the inverse, retrograde canon of music, as an application of symmetries. I was intrigued by this theme, so I asked the music teacher for help and she came to my rescue. We studied a path that started from music and resulted in the representation of symmetries on the Cartesian plane.

On this occasion too, the activity took place entirely during the mathematics classes: it was the mathematics teacher who asked the music teacher to collaborate on a specific topic and the mathematical content of the topic was the main aim of the activity. Figure 2 shows how the relationship between mathematics and the other disciplines could be conceptualized during this preliminary phase of the STEAM-Connect project.

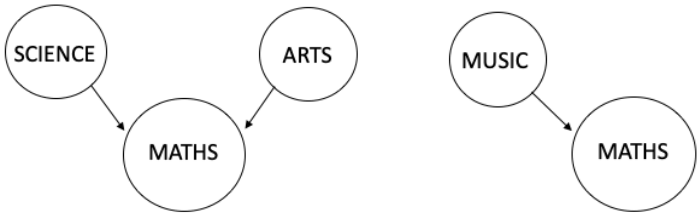


Figure 2. Relationships among disciplines at the beginning of the project.

In the following phases of the project, Paola and her colleagues built on these ideas on the topic of symmetries and translations, asking for the collaboration of the technology teacher to investigate the topic in the case of regular polygons. Paola declares that she was the promoter of the collaboration and that the mathematical content, in this phase, was still the aim of the whole path, even though the activity was designed to be carried out during the curricular classes of all the disciplines.

> Paola: Since the idea started from me and I started to involve the other colleagues, there was the mindset, even when talking about disciplinary contents, of reasoning in terms of mathematical contents.

Paola describes the path of the activity with the metaphoric image of the spiral, because it was designed to converge towards mathematics, in temporal and content terms.

> Paola: We made a sort of spiral, let's say, that is all the disciplines connected, even temporally, one after the other. […] So, the figure that best represents the path, in my opinion, is the spiral, which reaches the top, where mathematics is at the top.

Figure 3a shows how the relationship between mathematics and the other disciplines could be represented at this phase of the STEAM-Connect project. Mathematics is at the end of a path starting with arts, designed to introduce students to the mathematical content of symmetries and translations. Mathematics is, in a sense, the ideal *destination* of the spiral path.

Paola reports that, in the subsequent design cycles of the same activity, the growing connection among teachers enabled by the context fostered a new awareness that all disciplines could be integrated with equal dignity in the path. This awareness gradually changed the relationships between the different disciplines, resulting in other disciplines being no longer considered only at the service of mathematics, but all contributing to providing a wider understanding of a common topic with an interdisciplinary approach (Capone, 2022).

This evolution took place over several months in which the Italian teachers and researchers participating in the STEAM-Connect project met periodically to discuss the design work of STEAM activities and to compare the proposals of the different schools, thus promoting collective reflections.

> Paola: I believe that a path has been outlined in which everyone serves everyone. […] It's not just aimed at introducing mathematics content, I mean, it's not that the work that the technology colleague does, or what the art colleague does has only that purpose. It has a specific aim for each curricular discipline.

Paola also describes the role of mathematics in this advanced phase of the design process as a means of modelling what students discover in the other disciplines, as a lens that enables students to make sense of the complexity of a topic in an interdisciplinary way.

Paola: Mathematics ultimately re-read everything that the students did, in different contexts, with an eye to general modelling.

Figure 3b shows how the interdisciplinary relationship between mathematics and the other disciplines could be represented, in light of Paola's words, after several months of design work in the STEAM-Connect project. At the centre is the common topic of symmetries and translations connecting science, art, technology and music, with mathematics as a lens allowing to focus and model the connections. In this way, the design process as described by Paola reflects mathematics' "particular independence [which puts it] in a special position as regards interdisciplinary relationships" (Piaget, 1972, p. 127).

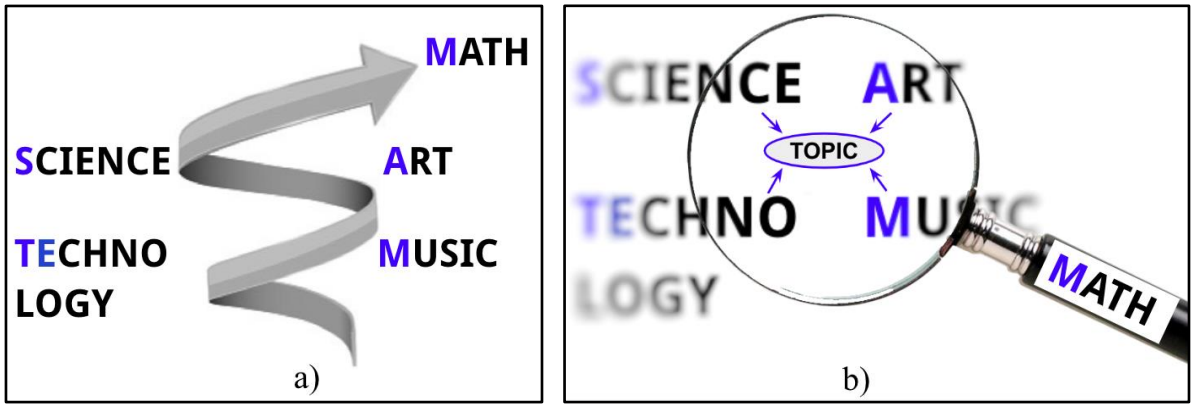


Figure 3. Evolution of the role of mathematics in the STEAM-Connect project.

According to Paola, the collaborative design work carried out in the context of the STEAM-Connect project fostered the awareness of the possibility – not often taken into consideration in Italian schools - to find common topics that could be addressed by different disciplines, creating a sort of "STEAM curriculum".

Paola: I believe that the main outcome of this collaborative design work is the awareness, raised in me and in my colleagues, of being able to find interconnections between the contents of different disciplines.

DISCUSSION

In this study, we conducted a *meta-disciplinary* (Williams & Roth, 2019) investigation, to understand the roles of the different disciplines in the design of a STEAM activity, particularly focusing on the evolution of the role of mathematics. The analysis of the case study revealed that the role of mathematics changed during the different design cycles of the activity. Indeed, during the first experiences of interdisciplinary activities conducted in Paola's school, the role of mathematics was predominant, with the mathematics teacher as the promoter of the collaboration with the art or the music teacher, who were at the service of the mathematical aim of the activity (Figure 2).

In the context of the STEAM-Connect project, as a consequence of the connection between teachers of different schools and researchers, reflecting together on the design of STEAM activities, the relationship between mathematics and the other disciplines started to evolve. The mathematical content of the topic of symmetries and translations

became the ultimate aim of a path, involving the other four disciplines in a "spiral" process leading towards it (Figure 3a). As the project progressed, the teachers collaborated on an increasingly equal basis to the design of an interdisciplinary STEAM activity with shared objectives, meaningful for all the involved disciplines (Figure 3b). The topic of symmetries and translations was perceived as a common object to all the disciplines, even though it was proposed by the mathematics teacher.

The different levels of connections favored by the STEAM-Connect project increased teachers' awareness of the possibility of collaboratively designing interdisciplinary activities on a shared topic. This awareness gradually moves them from a teaching paradigm focused on the knowledge products of the different disciplines to a teaching paradigm aimed at eliciting the processes that gave rise to the knowledge products and their mutual relations. A new awareness has also been reached of the modeling role of mathematics, which enables students to read natural phenomena or artistic expressions through a common lens.

Reflecting on how to connect different disciplines, each with its specific epistemology, in the design of an interdisciplinary "STEAM curriculum" could be both a developmental aim for the next phases of the project and a venue for future research.

ACKNOWLEDGMENTS

The research is partially funded by the STEAM-Connect Erasmus+ project (2021-1-LU01-KA220-SCH-000034433).

REFERENCES

Apostel, L., Berger G., & Briggs, A. (Eds.) (1972). *Interdisciplinarity: Problems of teaching and research in universities*. OECD Publishing.

Capone, R. (2022). Interdisciplinarity in Mathematics Education: From Semiotic to Educational Processes. *EURASIA Journal of Mathematics, Science and Technology Education*, 18(2). https://doi.org/10.29333/ejmste/11508

Cohen, L., Lawrence, M., & Keith, M. (2007). *Research Methods in Education. 6th Edition*. Taylor and Francis Group.

Doig, B., Williams, J., Swanson, D., Borromeo Ferri, R. & Drake, P. (Eds) (2019), *Interdisciplinary Mathematics Education. ICME-13 Monographs*. Springer.

Henriksen, D., Mehta, R., & Mehta, S. (2019). Design Thinking Gives STEAM to Teaching: A Framework That Breaks Disciplinary Boundaries. In M. S., Khine. & S., Areepattamannil (Eds.) *STEAM Education: Theory and Practice*. Springer. https://doi.org/10.1007/978-3-030-04003-1_4

Khine, M. S., Areepattamannil, S. (Eds.) (2019). *STEAM Education: Theory and Practice*. Springer. https://doi.org/10.1007/978-3-030-04003-1

Piaget, J. (1972). The epistemology of interdisciplinary relationships. In L. Apostel, G. Berger, & A. Briggs (Eds.), *Centre for Educational Research and Innovation, Interdisciplinarity: Problems of Teaching and Research in Universities* (pp. 127–139). OECD Publishing.

Pocalana, G., Robutti, O. (2022). Mathematics teacher educators work to foster an inquiry community. In Fernández, C., Llinares, S., Gutiérrez, A., & Planas, N. (Eds.), *Proceedings of the 45th Conference of the International Group for the Psychology of Mathematics Education (Vol. 3).* Alicante, Spain: PME.

Stohlmann, M. (2018). A vision for future work to focus on the "M" in integrated STEM. *School Science and Mathematics*, 1-10. https://doi.org/10.1111/ssm.12301

Watson, A. D. (2015). Design thinking for life. *Art Education*, 68(3), 12–18.

Williams, J. S., Roth, WM. (2019). Theoretical Perspectives on Interdisciplinary Mathematics Education. In Doig, B., Williams, J., Swanson, D., Borromeo Ferri, R., Drake, P. (Eds), *Interdisciplinary Mathematics Education. ICME-13 Monographs*. Springer.

Williams, J. S., & Wake, G. D. (2007). Metaphors and models in translation between college and workplace mathematics. *Educational Studies in Mathematics*, 64(3), 345–371.

UNIT STRUCTURES RARELY ARTICULATED: TEACHERS' EXPLANATIONS OF MEANINGS OF MULTIPLICATION

Susanne Prediger[1,2] and Anke Wischgoll[2]

[1] IPN Leibniz Institute for Science and Mathematics Education, Berlin, Germany

[2] TU Dortmund University

Explaining meanings is an important discourse practice for developing conceptual understanding of basic mathematical concepts such as multiplication. The study investigates how n = 102 teachers explained the meaning of multiplication, and how they reacted to a student's explanation. The analysis of teachers' written explanations reveals affordances and constraints: While the largest group of teachers referred only to symbolic repeated addition, more mathematics teachers than out-of-field teachers alluded to the key multiplicative structure of unitizing in their explanation, but both groups rarely articulated the unit structure explicitly. When reacting to a student's explanation with incomplete dot arrays, similar tendencies occur: Even when graphically referring to unit structures, only very few teachers offered an explicit language of unitizing. The findings show the need for professional development to enable all teachers to unpack the meaning of multiplication beyond repeated addition and to explicitly address meaning-related language for unitizing.

INTRODUCTION: INVESTIGATING TEACHERS' EXPLANATION

For developing conceptual understanding, the discourse practice of explaining meanings is crucial. It is usually conducted by teachers in instructional explanations or in interaction with students' explanations when teachers react to students' emerging explanations and leverage them (Leinhardt, 2001). However, instructional and co-constructed explanations can only enhance students' conceptual understanding under two empirically specified conditions: (a) it must *focus the crucial structure* underlying the concept in view (Wittwer & Renkl, 2008), (b) many students can only engage in explaining meanings focusing on structures when *targeted language* support *is provided* (Planas, 2021; Post & Prediger, 2022). Fulfilling these conditions requires considerable topic-specific (pedagogical) content knowledge (Leinhardt, 2001), so the concern arises that they might be particularly demanding for teachers without mathematics teaching certificates (called out-of-field teachers) who are increasingly employed in German schools. We compare out-of-field teachers' explanations with those of mathematics teachers. For this, we use a topic-specific approach for the topic of meanings of multiplication for natural numbers, with the following research question:

How do mathematics teachers and out-of-field teachers explain the meaning of multiplication, and how do they react to a student's incomplete explanation?

2023. 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. 59-74). PME 46.

THEORETICAL BACKGROUND

State of research on conceptual requirements: Meanings of multiplication

Many students have been shown to struggle with the meaning of multiplication, as it is more abstract than addition and subtraction, with different roles of both factors (Clarke & Kamii, 1996). The key meaning of multiplication is *counting in units*, i.e., 3 x 5 is to be interpreted as three units of fives (Götze & Baiker, 2021). For other countries, it was shown that many teachers still translate multiplication to *repeated addition* (3 x 5 means 5+5+5), converting between two symbolic representations (Askew, 2019); in spite of empirical evidence that interventions with repeated addition are less effective than those focusing more explicitly on unit structures (summarized by Clarke & Kamii, 1996). Jumps on number lines and dot arrays as graphical representations support students' meaning-making processes (Barmby & Milinkovic, 2011). Rather than only translating symbolic to graphical representations, *connecting* entails to *explain* how the structure of one representation is reflected in the other one (Post & Prediger, 2022); for multiplication, the unit structure is to be made explicit (Götze & Baiker, 2021).

State of research on linguistic requirements: Engaging in explaining meanings with language support of meaning-related phrases

In general, students' processes of meaning-making have been shown to be best promoted when they are engaged in discourse practices of explaining meanings (Moschkovich, 2015). Not only teachers' instructional explanation offers or constraints learning opportunities (Lachner & Nückles, 2016), but also the interactively constructed explanations in which teachers react to and leverage students' emergent ideas (Leinhardt, 2001). However, not all students can immediately contribute to these discourse practices (Moschkovich, 2015), so teachers are requested to provide language support, with phrases that can be used not only to articulate procedures, but also the mathematical structures underlying the meaning of a concept; these phrases are *termed meaning-related phrases* (Post & Prediger, 2022). Case studies found teachers' struggle with identifying and promoting meaning-related phrases (ibid.; Planas, 2021).

Substantiating these considerations for the topic of meaning-making for multiplication, it is not enough to say, e.g., "the dot array matches 3 x 5 because here is 3 and here is 5". Instead, the unit structure needs to be articulated, e.g., by "the dot array consists of three fives" or "three rows of five each". Götze & Baiker (2021) showed that teaching with language offers of such meaning-related phrases results in significantly higher learning gains than with the same tasks and graphical representations, but without meaning-related phrases articulated. When studying the ways teachers explain multiplication, it is therefore crucial to also assess this linguistic requirement, the explicitness in which they articulate the unit structure by such kind of meaning-related phrases.

State of research on teaching requirements: Teachers' explanation of meaning and potential challenges for out-of-field teachers

Teachers' explanations of meanings have been studied for various mathematical concepts, e.g., optimization problems (Lachner & Nückles, 2016), or multiplication of fractions (Shure et al., 2022). Both studies reveal that many teachers tend to prioritize reports of procedures before explanations of meanings, and Lachner & Nückles (2016) identified that teachers with lower content knowledge have a higher risk for this prioritization. Shure et al. (2022) show that prospective teachers rarely meet the linguistic requirements of offering meaning-related phrases in the case of fractions. With respect to multiplication with natural numbers, Barmby & Milinkovich (2011) showed that teachers have heterogeneous abilities to choose adequate representations meeting the conceptual requirements, but their first insights focus only on conceptual requirements, but not on linguistic requirements, similar to the teachers observed by Askew (2019). For all topics, little is known about how teachers' instructional explanations are associated with their reaction to students' incomplete explanations, although this is the first step for interactively establishing collective explanations.

Given the repeated findings that (pedagogical) content knowledge can heavily influence teachers' explanations (Leinhardt, 2001; Lachner & Nückles, 2016), a particular group of teachers raises our interest which is currently growing due to mathematics teacher shortage: About 20-30% of the teachers who teach mathematics in German middle schools hold no mathematics teaching certificate. Several international studies indicate that students learn less from out-of-field teachers (Hobbs & Törner, 2019), but so far, the mechanisms are not understood in detail. In the context of explaining meanings for multiplication, we *hypothesize that out-of-field teachers might provide explanations that meet less the conceptual and linguistic requirements than mathematics teachers*. By testing this hypothesis, we intend to contribute to explaining the mechanisms underlying the findings on out-of-field teachers' lower effectiveness.

METHODS

Methods of data gathering

Sample. The sample consisted of German teachers (n = 102) in the beginning of their first session of a volunteer professional development (PD) online course on teaching conceptual understanding of basic concepts. 65 reported to hold a mathematics teaching certificate, 29 to be out-of-field teacher, 8 did not disclose their certificate. The teachers had 0-36 years of experience in mathematics teaching (median of 4 years).

Data collection and participant consent. A LimeSurvey was conducted in the first 15 minutes of the PD course. For the participants, it was framed as an activity to first think individually about key questions of the PD that were later discussed in small groups. The research purpose was explained, and participants' consent was collected.

Items on explaining and reacting to student explanation. This paper focuses on two open items on meaning-making for multiplication (Figure 1): Item 1 asks the teachers to give an own explanation, it is slightly scaffolded by an uncommented image (of six dice of four dots each) which only few teachers used in their explanation. Item 2 captures the first step of interactively establishing explanations, asking the teacher to react to a student's explanation in an incomplete dot array that clearly hints to a well-known superficial idea of translating representations without reference to the unit structure.

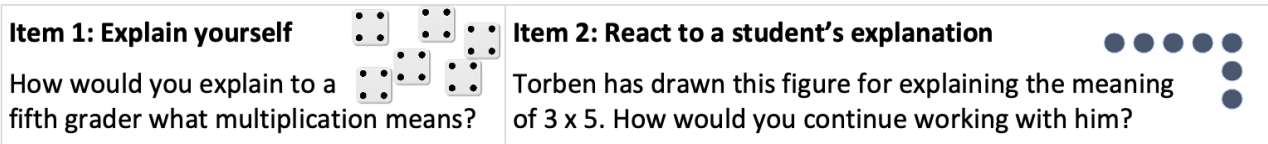


Figure 1. Two items for teachers in view of this study

Methods of data analysis

The teachers' written responses were analyzed for both items in a deductive-inductive code formation process (Mayring, 2015). In *Step 1*, we started with codes known from the literature on conceptual and linguistic requirements and teachers' typical forms of explanation for multiplication. In *Step 2*, these deductively given codes were then refined and extended inductively to capture the relevant phenomena in the data. For saving space, the resulting coding schemes with anchoring examples will be presented together with the frequencies in Tables 1/2 in the next section. In *Step 3*, all 102 written answers were rated by two coders who reached an interrater reliability of Cohens κ = .83. In *Step 4*, the frequencies of different explanations and reactions were determined for the whole sample and compared between the subsamples of mathematics teachers and out-of-field teachers. The hypothesis that out-of-field teachers meet less of the conceptual and linguistic requirements than mathematics teachers was tested by Fisher's exact tests in place of χ^2 tests (due to small sample sizes) on the 5 % level, as well as the associations identified in contingency data of Item 1 and 2.

RESULTS

Teachers' own explanations of the meaning of multiplication and differences between mathematics and out-of-field teachers

Table 1 presents all codes developed to capture the teachers' written explanations for Item 1 (from Figure 1) and anchoring examples. The third column lists the frequencies of codes for explanations in the whole sample of all 102 teachers. It reveals that 20% of the teachers gave no explanation or none that met conceptual and linguistic requirements. 65% gave an explanation that is conceptually acceptable, but not yet articulated in a concise and explicit language (the linguistic requirement), among them 31% who mentioned only repeated addition. 15% of the teachers provided explanations that met

conceptual and linguistic requirements, with 13% using a condensed phrase that has been identified as preferable in earlier classroom research (Götze & Baiker, 2021).

Table 1. Frequency of codes for teachers' explanations (Item 1)

Code	Anchoring examples	Percent of explanations of …		
		all teachers (n = 102)	**out-of-field teachers** (n= 29)	**mathematics teachers** (n= 65)
Not meeting conceptual and linguistic requirements		***20 (20%)***	***10 (35%)***	***5 (8%)***
Missings	[No answer given]	11	4	3
No reference to meaning	"Multiplication means times."	9	6	2
Meeting conceptual but not linguistic requirements		***66 (65%)***	***15 (52%)***	***49 (75%)***
Only repeated addition	"6 x 4 is 4+4+4+4+4+4" "Addition done several times."	32	6	26
Only duplication	"Duplication of numbers."	5	2	2
Unitizing, partially articulated	"You count how many [packages] and how many [in each] of them." or "4 each"	29	7	21
Meeting conceptual and linguistic requirements		***16 (15%)***	***4 (14%)***	***11 (17%)***
Unitizing, articulated in 2 phrases	"Six dice. And four dots on each them." or "6 packages, in the package there are four."	3	1	2
Unitizing in one phrase	"Three groups of four each." "Three of the fours."	13	3	9

The last two columns compare frequencies for mathematics and out-of-field teachers. Out-of-field teachers gave much more often no explanation or none that fulfills the requirements than the mathematics teachers (35% vs. 8%). Significantly less often they provided a conceptually acceptable answer, yet without explicitly articulating the unitizing structures (52% vs. 75%), with repeated addition occurring half as much (20% vs. 40%). The 2x3 Fisher exact test confirms that as hypothesized, out-of-field teachers met the conceptual requirements with highly significantly lower frequency ($p < .01$), while no substantial differences were found for linguistic requirements (14% vs. 17%).

Teachers' reactions to a student's explanation and differences between mathematics and out-of-field teachers

Table 2 documents the codes developed to capture the teachers' written reaction to Torben's wrong explanations for Item 2 in Figure 1, with anchoring examples. Only two reactions falsely evaluated Torben's answer as correct. 15 teachers reported to first ask back (which can indeed be important). So, the total of 31% of reactions not fulfilling conceptual and linguistic requirements contains these 15% asking back which might meet the requirements in the next steps. 59% of the teachers suggested to react by pointing the student to the meaning of multiplication in dot arrays (36% of them without unitizing, only focusing repeated addition or the total of 15 dots), and without providing any language support for articulating the unit structure. Only 10% articulated

the unit structures explicitly and thereby fulfilled the conceptual and linguistic requirements in their reaction (even less than in the own explanation in Item 1).

Table 2. Frequency of codes for teachers' reaction to Torben's explanations (Item 2)

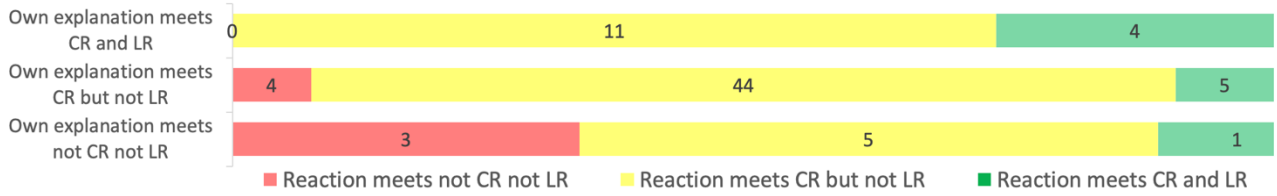
Code	Anchoring examples	Percent of explanations of …		
		all teachers (n = 102)	**out-of-field teachers** (n= 29)	**mathematics teachers** (n= 65)
Not meeting conceptual and linguistic requirements		***32 (31%)***	***13 (45%)***	***13 (20%)***
Missing	[No answer given]	15	8	2
Falsely evaluated as correct	"Correct"	2	0	1
Asking back without further plan	"I need to ask back what he thinks about it."	15	5	10
Meeting conceptual but not linguistic requirements		***60 (59%)***	***12 (41%)***	***46 (71%)***
Repeated addition	"Make him see the repeated addition of rows."	3	1	2
Complete dot array without unitizing	"You have to draw all dots because we want to see 15."	33	5	27
Unitizing, incompletely articulated	"I want him to see the 3 complete rows."	24	6	17
Meeting conceptual and linguistic requirements		***10 (10%)***	***4 (14%)***	***6 (9%)***
Unitizing explicitly articulated	"I ask for showing me 3 fives." or "Three rows of fives."	10	4	6

Again, significant differences can be found between mathematics and out-of-field teachers in these reactions: The 2x3 Fisher exact test confirms the hypothesis that out-of-field teachers met significantly less the conceptual requirements than mathematics teachers ($p < .05$), whereas the linguistic requirements were even met minimally better.

Associations between own explanation and reaction to student explanation

Figure 2 visualizes the contingency data how teachers' own explanations (in Item 1) and their reactions to students' explanations (in Item 2) are associated. The bars document the rating of teachers' own explanation in Item 1: 15 explanations in the upper bar met both, conceptual and linguistic requirements, 53 in the medium bar met conceptual but not linguistic requirements, and 9 in the lowest bar meet none of the requirements (missings and codes "asking back" were excluded, so the total is 77).

Figure 2. Associations of teachers' own explanation to reactions to explanations



The colors in the bars indicate the connection to the teachers' suggested reactions in Item 2, with a strong dominance of yellow, i.e., meeting conceptual but not linguistic

requirements: Out of 9 teachers whose explanation in Item 1 met no requirements, 33% (3 of 9) suggested reactions not meeting the conceptual and linguistic requirements in Item 2. This percentage is higher than 4 of 53 for those with medium explanations and 0 with top explanations. In contrast, 26% (44 of 15) of those with top explanations also met both requirements in their reaction in Item 2. The 3x3 Fisher exact test shows that these associations between own explanations and reactions are significant ($p < .05$).

DISCUSSION

Whereas other studies have analyzed teachers' explanations in real classroom situations (Leinhardt, 2001; Planas, 2021; for multiplication Askew, 2019), with all complexities and pressure to instantaneous reaction, this study investigated explanations and reactions written in an action-relieved situation in the beginning of a PD session. Table 1 reveals that 80% of the teachers offered explanations that largely met conceptual requirements, including 31% repeated addition, which was evaluated as possibly bearing an adequate conceptual structure (equal groups being united), but not a suitable language to articulate the unit structures explicitly, as they are either only formal or focusing on joining, but not on generating new units. In contrast, only 15% of all teachers' explanations also met linguistic requirements by explicitly articulating the unit structures. The percentages are even lower in Table 2 for teachers' reactions to a false student explanation (69% meeting conceptual requirements, 10% for linguistic requirements). The contingency data in Figure 2 reveals that teachers who gave full explanations of the meaning of multiplication (that adequately address and explicitly articulate the multiplicative structure) were more likely to react to students' incomplete ideas in conceptually adequate ways, but not all of them provided good language support for students to articulate the multiplicative unit structure explicitly.

The hypothesis that out-of-field teachers meet the conceptual requirements significantly less frequently has been confirmed (and resonates with often documented conceptual restrictions in out-of-field teaching, Hobbs & Törner, 2019). But the hypothesis was rejected for linguistic requirements: Regarding explicit language support, mathematics and out-of-field teachers show similarly low frequencies.

These findings must be interpreted with caution, given the *methodological limitations* of the study (only two items and a relatively small sample size for group comparisons).

But already now, the findings show a *strong need for professional development on conceptual and linguistic requirements* for more helpful explanations, in particular enabling teachers to give explicit language support by meaning-related phrases (Planas, 2021; Götze & Baiker, 2021; Post & Prediger, 2022). During the PME conference, we hope to be able to present also pre-post data of the evaluation study of the PD and se changes of teachers' explanations through the PD.

Acknowledgement. This study is a first step of the MuM video PD study (BMBF grant No. 01JD2001A to S. Prediger by the National Ministry of Education and Research).

REFERENCES

Askew, M. (2019). Mediating primary mathematics: measuring the extent of teaching for connections and generality in the context of whole number arithmetic. *ZDM – Mathematics Education, 51*(1), 213–226.

Barmby, P., & Milinkovic, J. (2011). Pre-service teachers' use of visual representations of multiplication. In B. Ubuz (Ed.), *Proceedings of 35th PME* (Vol. 2, pp. 105–113). PME.

Clark, F. B., & Kamii, C. (1996). Identification of multiplicative thinking in children in grades 1-5. *Journal for Research in Mathematics Education, 27*(1), 41–51.

Götze, D., & Baiker, A. (2021). Language-responsive support for multiplicative thinking as unitizing: results of an intervention study in the second grade. *ZDM – Mathematics Education, 53*(2), 263–275.

Hobbs, L., & Törner, G. (Eds.). (2019). Examining the phenomenon of "teaching out-of-field": International perspectives on teaching as a non-specialist. Singapore.

Lachner, A., & Nückles, M. (2016). Tell me why! Content knowledge predicts process-orientation of math researchers' and math teachers' explanations. *Instructional Science, 44*(3), 221–242.

Leinhardt, G. (2001). Instructional explanations. In V. Richardson (Ed.), *Handbook of research on teaching* (pp. 333–357). American Educational Research Association.

Mayring, P. (2015). Qualitative content analysis. In A. Bikner-Ahsbahs, C. Knipping, & N. Presmeg (Eds.), *Approaches to qualitative research in mathematics education: Examples of methodology and methods* (pp. 365-380). Springer.

Moschkovich, J. (2015). Academic literacy in mathematics for English Learners. *The Journal of Mathematical Behavior, 40*(A), 43–62.

Planas, N. (2021). How specific can language as resource become for the teaching of algebraic concepts? *ZDM – Mathematics Education, 53*(2), 277–288.

Post, M., & Prediger, S. (2022, online first). Teaching practices for unfolding information and connecting multiple representations. *Mathematics Education Research Journal*. https://doi.org/10.1007/s13394-022-00431-z

Shure, V., Rösken-Winter, B., & Lehmann, M. (2022). How pre-service primary teachers support academic literacy in mathematics in a scriptwriting task encompassing fraction multiplication and division. *The Journal of Mathematical Behavior, 65*:100916, 1–16.

Wittwer, J., & Renkl, A. (2008). Why instructional explanations often do not work. *Educational Psychologist, 43*(1), 49–64

STUDYING THE ROLE OF PSEUDO-OBJECTS IN PROOF BY CONTRADICTION

Kostas Probonas and Giorgos Psycharis

National and Kapodistrian University of Athens

This paper reports research focusing on how pseudo-objects (i.e. geometrical figures in a dynamic geometry environment containing contradictory properties) may enhance upper secondary students' learning of proof by contradiction. Under the lens of the cognitive unity of Theorems and Theory of Semiotic Mediation we explore the semiotic potential of pseudo-objects in mediating the proving process from the conjecturing phase to the final proof. We analyzed the work of three pairs of students during their engagement with non-constructability tasks favoring the occurrence of pseudo-objects. The results show that pseudo-objects contributed differently to students' argumentation and proof. This indicates the complexity of addressing the semiotic potential of pseudo-objects in students' learning of proof by contradiction.

INTRODUCTION AND THEORETICAL FRAMEWORK

Many research studies have focused on analyzing how dynamic geometry environments (DGE) mediate the transition from conjecture to proof through the development of argumentation. In most of these studies, it is stated that there is a cognitive gap between the formal proof and the empirical argumentation developed by students who aim to construct a proof (Mariotti, 2019). Boero et al. (1996) introduced the idea of cognitive unity of theorems to describe the continuity and connection between the arguments developed in the construction and validation of a conjecture and those used in the construction of the proof. A proof may be more easily accessible to the student, if it is related to a previous activity of argumentation aimed at supporting the production of a conjecture.

A specific type of proof is indirect proof, such as proof by contradiction, which can be defined as 'the proof of a statement, whose premise includes the negation of the conclusion' (Baccaglini-Frank et al., 2018). If a statement S can be expressed as an implication $p \rightarrow q$, a proof by contradiction is the direct proof of the proposition S*: $p \cap \neg q \rightarrow r \cap r \neg$, where r is any proposition (Antonini & Mariotti, 2008). Accordingly, an indirect argument is an argument in the form '…if it were not so, it would happen that…' (ibid). Over the last years there is an increased research attention to the role that DGS may play in proof by contradiction especially through the use of pseudo-objects (Leung & Lopez, 2002). A pseudo-object is defined as 'a geometrical figure associated to another geometrical figure either by construction or by projected-perception in such a way that it contains properties that are contradictory in the Euclidean theory (Baccaglini-Frank et al., 2013, p. 65). Recent studies have pointed out the specific contribution that DGE have in supporting students' indirect argumentation that can lead to proof by contradiction. To perceive a pseudo-object,

2023. 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. 67-74). PME 46.

one must perceive in a figure contradictory geometric properties in the context of Euclidean geometry. Such properties in a DGE can result from the conflict of direct and indirect invariants. This perception could occur when the figure is degenerated due to the co-existence of the conflicting properties (Baccaglini-Frank et al., 2018). The direct invariants are determined by the geometric relations defined by the commands used to construct the figure (robust constructions) and the indirect invariants result from the consequences of the construction in the context of Euclidean geometry (Mariotti, 2019). Baccaglini-Frank et al. (2013; 2018) highlighted the important role of pseudo-objects in students' reasoning during the process of proof by contradiction. Specifically, they confirmed that its occurrence in a DGE is a decisive factor that influences the development of indirect argumentation and proof.

Our study is also informed by the Theory of Semiotic Mediation (TSM) (Bussi & Mariotti, 2008) that recognizes the central role of the semiotic process (i.e. transformation of signs) in teaching and learning. Mariotti (2019) explored the role of dynamic geometry tools as tools of semiotic mediation, the process of internalizing these tools and their use in developing geometric thinking capable of bridging the gap between argumentation and proof. The logical dependence of robustly constructed properties with their consequences (invariants) can be related to a Theorem. A Theorem is a system of elements where a proof can be conceived when there is a statement to which it provides validation, but also when there is a theory (e.g., Euclidean Geometry) within which the validation makes sense (Mariotti, 2019). Thus, each construction in a DGE has a counterpart in the statement of a Theorem. Mariotti (2019) also highlighted the semiotic potential of a DGE in helping students develop mathematical meanings related to a Theorem and the semiotic potential of a pseudo-object, which is related to the mathematical meaning of indirect proof.

The purpose of the present research is to study the contribution that DGE have in supporting students' indirect argumentation that can lead to proof by contradiction. Under the lens of the cognitive unity of Theorems and TSM, we aim to explore the semiotic potential of the pseudo-object in mediating the proving process from the conjecturing phase to its final product.

METHODOLOGY

Participants and data collection

Data were collected by the first author through three interventions of three teaching hours each in a private school. The participants were 3 pairs of students, two pairs from grade 11th and one from grade 10th (4 girls, 2 boys). Each intervention was preceded by a students' familiarization with the The Geometer's Sketchpad software lasting 2 teaching hours as the students had no previous experience with it. Data collected through the screen and audio recorders that students had on their laptops as well as a camera that recorded their movements towards the screen. Data were fully transcribed for the analysis.

Research tools

Students were given three open, conjecturing, non-constructability tasks in succession, which were preceded by a task where the figure was constructable. All three problems given to the students were constructed by the first author according to Baccaglini-Frank et al.'s (2017) design principles with the aim to facilitate students' production of indirect argumentation and indirect proof. In the present study, we utilized the type of tasks given in the form 'Is it possible to construct a figure of type X with properties Y_1, Y_2, … Y_n? If so, construct it robustly. If not, explain why not.'. Our a-priori analysis of the tasks confirmed this potential, in addition to the potentiality of the figures to turn into a pseudo-object.

Task 1: Can you construct a right triangle whose exterior bisector of an angle is parallel to one of its sides? If yes, describe the construction steps. If not, explain why.

Task 2: Can you construct a rectangular parallelogram ABCD such that the bisector of the angle with sides AD and the diagonal AC and the bisector of the angle with sides BC and the diagonal BD are perpendicular? If yes, describe the construction steps. If not, explain why.

Task 3: Construct an isosceles triangle ABC with AB=AC and the circle with center A and radius $r \leq AB$. Let K, L be the points of intersection of the circle with AB and AC respectively. Construct CK, BL (T the point of intersection). Can CK and BL be perpendicular and simultaneously be tangent to the circle at points K,L? If yes, describe the construction steps. If not, explain why.

METHOD OF ANALYSIS

The analysis was carried out in two steps following a broadly grounded approach (Charmaz, 2014). In the first step, we selected the transcripts concerning the students' interaction with pseudo-objects and from them we extracted dialogical episodes following the development/evolution of students' argumentation (conjectures, kind of arguments, contradictions, proofs). For each dialogical episode (a) we coded the students' actions (e.g. dragging) while working with the DGE, the feedback that students received and its source and (b) we provided analytical comments according to our theoretical framework (cognitive unity, TSM). In the second step, we aimed to make explicit the constituent elements of students' argumentation in terms of geometric properties and relationships (invariants). For this, we constructed tables for each one of the identified contradictions coding in columns: direct invariants; induced condition (by the student); indirect invariants; and contradictions. In the third step, constant comparison between the coded dialogical episodes and tables (first and second steps) allowed us to categorize the students' approaches in terms of the formulation of the conjecture, the development of direct/indirect argumentation and the role of the pseudo-object in this proving process.

RESULTS

Two categories emerged from the data analysis. In the first category, the students used productively the feedback of the pseudo-object to develop their argumentation in connection with geometric theory and constructed a valid proof. In the second category, the students utilized the pseudo-object mostly for visual evidence. This limited their argumentation to visual and empirical arguments. Below, we present two illustrative episodes from the work of two groups of students to highlight each category. The analysis is accompanied by two tables (tables 1,2) that represent the geometric properties and their relationships that constitute the students' perception of the pseudo-object during the proving process (from conjecture to proof).

Constructing a valid proof: The case of Clio and Silia

In the episode of students Clio and Silia in task 1, the students constructed the right triangle (A1, table 1) and the external bisector of C (A2) robustly and after a short exploration they tried, through dragging the point B, to induce that the external bisector is parallel with AB (P1). The triangle was degenerated into a straight line (figure 1) and that was perceived by students as a result of the conflict of direct (A1⋂A2) and indirect invariants (P1). The initial contradiction (A1⋂A4) arose as visual feedback of the degeneration of the triangle (figure 1b) and the figure turned into a pseudo-object. Then the students made a conjecture concerning the non-constructability of the figure.

Clio: (They become parallel) only if AC...

Silia: ... is coincide with BC. But ... that can't be done. Because it won't be a triangle after that.

By dragging point B close to point A, students explored further the dynamic changes in the figure with the aim to justify its degeneration. In this process, they developed indirect argumentation to support their conjecture about the non-constructability of the figure. The arguments were in the form 'if they (AB, CD) become parallel then it will happen...'. Eventually, the students perceived the indirect invariant that the external bisector bisects an angle of 180° (P2), which contradicts with the geometric property that each external angle of a triangle is less than 180º (L1). This contradiction was also visualized on the screen (figure 1b).

Silia: It should be 90º (to be parallel)

Clio: The whole angle should be straight. [*Points to external angle*]

Silia: It should be vertical (the line bisector) … but the external C is not 180º.

The students seemed confident about the validity of their arguments and the researcher asked them to construct a proof in a piece of paper. For about ten minutes the students were attempting to complete the proof on the paper and at the same time they continued to interact with the DGE figure (dragging point B close to point A). Also, they pointed to angles and sides with their fingers on the screen and simultaneously referred to geometric properties of the figure in each step of their proof construction.

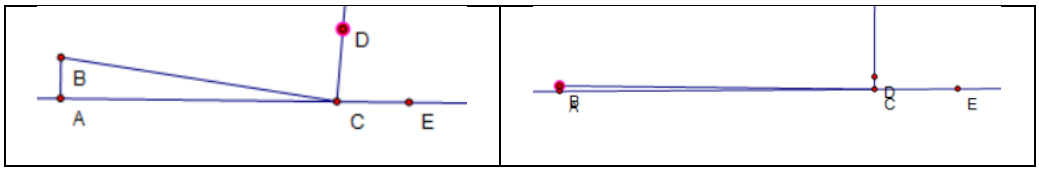


Figure 1a,b: Figure and degenerated figure (task 1)

Finally, they completed their proof on the paper through the use of symbols and mathematical relationships. The final outcome of their work was influenced by the feedback of the pseudo-object and the visualization of the contradictions it offered. During the analysis, we (the authors) were able to describe this influence by interrelating the students' work on DGE, their gestures pointing to the screen, their oral expressions (e.g. about geometrical properties stemming from the invariants) and the written parts of the proof on the paper. For instance, the indirect invariant 'the external bisector bisects an angle of 180°' (P2) was matched with the expression BCD+BCA=DCE while the direct invariant 'CD is an external bisector' (A2) was matched with BCD=DCE.

Clio: [*Showing on the screen*] If the external bisector bisects an angle of 180°… Ah, this (BCD) plus this (BCA) should equal this (DCE), but it can't because this (BCD) is equal to this (DCE).

In this episode the visual feedback that the DGE provided through the degeneration of the figure allowed the students to formulate the conjecture. Their interpretation of degeneration was supported by dragging that allowed them to reproduce the degeneration, explore its conditions and relate it with geometrical properties.

Direct invariants	Induced condition	Indirect invariants	Contradiction
A1∩A2	P1	A3(⇒A4)	
A1∩A2	P1	P2(⇒A4)	
Final proof in paper			
A1∩A2	P1	P2: BCD+BCA=DCE A2 : BCD=DCE	(P2∩A2:BCA=0) ∩ L1
A1: ABC is a right triangle with A=90°, A2: the line CD is the external bisector, A3: AB coincides with BC, A4: ABC is not a triangle (degenerates into a straight line) P1: external bisector is parallel to AB, P2: the external bisector bisects an angle of 180° (the external angle is 180°) L1: each external angle of a triangle <180o			

Table 1: Analysis of students' perception of contradictions (task 1)

This interplay of visualization and geometric theory mediated by the dragging tool, led students' argumentation to evolve and helped them to explain mathematically the non-constructability of the figure. Their arguments developed in the conjecturing phase were reorganized and transformed into a proof by contradiction under the lens of the Euclidean geometry. Thus, the students related the DGE figure to a Theorem and through this they achieved cognitive unity.

Constructing an invalid proof: The case of Fani and Valia

This episode is extracted from the work of Fani and Valia on task 2. The students had constructed robustly the requested figure (A1, A2 table 2) and they started to drag its draggable points so as to achieve the verticality of the bisectors (A1). The figure degenerated into a straight line (A3, figure 2b). The degeneration caused uncertainty to students about the constructability of the figure and challenged them to provide explanations.

Valia: To be perpendicular (the bisectors), ABCD becomes a straight line. But is this (ABCD) a parallelogram?

Fani: So, when it becomes a straight line, the angle becomes 180°… right?

Valia: Zero degrees, isn't it? So, they bisect an angle (DAC) of 0° perhaps…

The students realized that the size of the bisected angle was a consequence (indirect invariant) of the condition they were trying to impose (bisectors being perpendicular). So, they used the angle measure tool to measure it. The feedback of the measure (90°) allowed the students to discern the perpendicularity of the bisected angle DAC (P2). This was perceived as a contradiction (P2∩L1) and led them to state the conjecture of non-constructability.

Valia: So, to become 90° (AEB)

Fani: DAC will also become 90°

Valia: That means that it is not possible…

Fani: A part of DAC can't be 90° by itself!

The pseudo-object emerged as the students perceived the contradictory geometric properties of an angle being 90 degrees that had to be less.

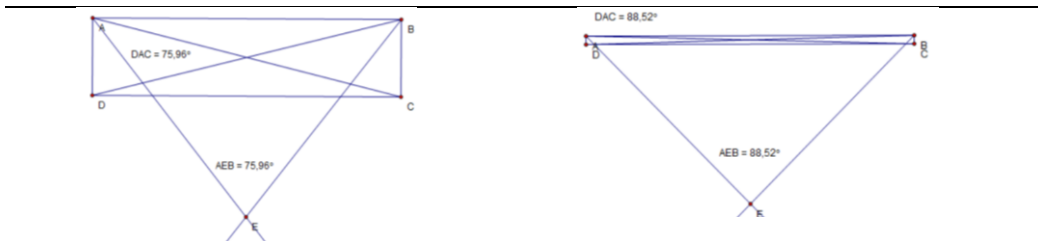


Figure 2 a,b: Figure and degenerated figure (task 2)

The indirect argument had as a source the visual feedback of pseudo-object and the angle measure tool. However, the contradictory geometric properties of the pseudo-object were perceived by the students visually and empirically. The argument 'when the bisectors become perpendicular, the bisected angle DAC becomes 90 degrees' was not explained in connection to Euclidean Geometry.

Direct invariants	Induced condition	Indirect invariants	Contradiction
A1∩A2	P1	A3	
A1∩A2	P1	P2,A3 (So, A4)	

Final proof in paper			
A1∩A2	P1	P2,L2 (So A4)	P2∩L1 and A5∩L2
A1: ABCD is a rectangular parallelogram, A2: Lines AE and BE are the bisectors of angles DAC and DBC, A3: Points, sides and diagonals coincide, A4: ABC is not a rectangular parallelogram (it degenerates into a straight line), A5: The sum of the angles of the parallelogram is 360°, P1: The angle between the two bisectors is 90°, P2: The bisected angle is 90°, L1: Bisected angle is <90° (sub-angle), L2: The sum of the angles of the parallelogram is >360°			

Table 2: Analysis of students' perception of contradictions (task 2)

The final product of the students' work was an invalid proof (table 2) due to the fact that it was based on the above mentioned empirical argument. Their proof had an explanatory function as it persuaded them for the non-constructability of the figure, but their conjecture was not validated within geometric theory. This episode indicates the semiotic potential of the pseudo-object with respect to the mathematical meaning of the contradiction. Although the students did not provide a mathematically valid proof, their interaction with the pseudo-object provided a basis for developing a conjecture through indirect argumentation and conceptualizing the contradiction at least in the DGE figure. The episode indicates also the diverse forms of cognitive unity that can be actualized in the process of rebuilding a conjecture into a mathematical theorem.

CONCLUSION

The results of the present study highlight how the DGE can mediate the proving process in open non-constructability tasks designed to promote the occurrence of a pseudo-object during students' exploration. The degeneration of the DGE figure provoked students' engagement in exploring further its properties as well as in providing explanations through indirect argumentation. The analyzed cases of students reveal that the pseudo-object contributed differently to students' argumentation and proof. In the first case, it facilitated proof validation through connections to geometric theory. In the second case, it facilitated the construction of an invalid and explanatory proof promoted by the visual feedback. Our findings also indicate the critical role of pseudo-objects in the final product of proof. In both analyzed cases, the (valid and invalid) proof developed by the students on the paper appeared to be strongly related to the arguments they developed during their interaction with pseudo-objects. In conclusion, our study shows that pseudo-objects do not concern only the perception of a pair of contradictory properties, but they can foster a dynamic process of investigating, looking for geometrical properties, explaining emerging (visual or theoretical) conflicts and developing argumentation within geometric theory. Also, our study reveals the diverse ways by which the semiotic potential of pseudo-object can be actualized and contribute to the cognitive unity of Theorems in proof by contradiction. This brings to the fore the need for further and systematic research in the role of pseudo-objects in students' learning of proof by contradiction.

REFERENCES

Antonini, S., & Mariotti, M. A. (2008). Indirect proof: what is specific to this way of proving?. *ZDM*, *40*(3), 401-412.

Baccaglini-Frank, A., Antonini, S., Leung, A., & Mariotti, M. A. (2013). Reasoning by contradiction in dynamic geometry. *PNA*, *7*(2), 63-73.

Baccaglini-Frank, A., Antonini, S., Leung, A., & Mariotti, M. A. (2017). Designing non-constructability tasks in a dynamic geometry environment. In A. Leung & A. Baccaglini-Frank (Eds.), *Digital technologies in designing mathematics education tasks—Potential and pitfalls* (pp. 99–120). New York: Springer

Baccaglini-Frank, A., Antonini, S., Leung, A., & Mariotti, M. A. (2018). From pseudo-objects in dynamic explorations to proof by contradiction. *Digital Experiences in Mathematics Education*, *4*(2), 87-109.

Boero, P., Garuti, R., & Mariotti, M. A. (1996). Some dynamic mental processes underlying producing and proving conjectures. In L. Puig & A. Gutiérrez (Eds.), *Proceedings of the 20th Conference of the International Group for the Psychology of Mathematics Education (PME 20)*, (Vol. 2, pp. 121-128). University of Valencia, Valencia.

Bartolini Bussi, M. G., & Mariotti, M. A. (2008). Semiotic mediation in the mathematics classroom: Artifacts and signs after a Vygotskian perspective. In L. D. English (Ed.), *Handbook of international research in mathematics education* (pp. 750–787). Mahwah, NJ: LEA.

Charmaz, K. (2014). *Constructing grounded theory*. London, England: SAGE.

Leung, A., & Lopez-Real, F. (2002). Theorem justification and acquisition in dynamic geometry: A case of proof by contradiction. *International Journal of Computers for Mathematical Learning*, *7*(2), 145-165.

Mariotti, M.A. (2019). The Contribution of Information and Communication Technology to the Teaching of Proof. In: G. Hanna, D. Reid, & M. de Villiers (Eds), *Proof Technology in Mathematics Research and Teaching*. *Mathematics Education in the Digital Era*, vol 14. Springer, Cham.

LEVELS OF MATHEMATICAL KNOWLEDGE IN LINEAR ALGEBRA FOR ENTERING UNIVERSITY

Kolja Pustelnik[1], Stefanie Rach[1], Daniel Sommerhoff[2], and Stefan Ufer[3]

[1]Otto-von-Guericke-University Magdeburg, Germany
[2]IPN – Leibniz Institute for Science and Mathematics Education, Kiel, Germany
[3]Ludwig-Maximilians-University Munich, Germany

Transitioning from school to university represents a challenge for most mathematics students, which is often connected to different characteristics of mathematics at both institutions. Prior mathematical knowledge has been shown to play an important role for success in the first year of university studies. Based on an IRT-approach, we investigate a model for prior knowledge in Linear Algebra for the transition from school to university. The model comprises four levels of mathematical demand and is based on an analogous model for Analysis. It is grounded on the distinction between conceptual and procedural knowledge and on changes of representations. The model allows to describe students' prior knowledge and investigate in further research, which kind of knowledge is important for the transition to university.

INTRODUCTION

The transition from school to university represents a challenge for many mathematics students. In Germany, many students with a major in mathematics drop out (Heublein, 2014). While the reasons for this dropout are manifold, several studies showed that students' prior mathematical knowledge is an important predictor for success at the transition from school to university (Hailikari et al., 2008; Halverscheid & Pustelnik, 2013). However, to understand the transition and as a basis for developing means to support students, *more is better* findings are insufficient, and it is vital to characterize levels of mathematical knowledge that are helpful for a successful transition.

For the domain of Analysis, the KUM-A model (Rach & Ufer, 2020) describes levels of mathematical knowledge considered relevant for a successful transition to university. While the model consists of four levels of demand, empirical data indicates that the third level distinguishes between students who master their first Analysis course and those who do not. Items on this level require a deeper understanding beyond well-known procedures and representations.

In this article, we present a corresponding model for Linear Algebra based on a test with 25 items. Therefore, we describe four levels of knowledge and investigate the item assignment to a level based on the item's empirical difficulty. Using student solutions for two items, we illustrate the validity of the model.

2023. 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. 75-82). PME 46.

THEORETICAL BACKGROUND

Mathematics at the transition from school to university

At the beginning of university studies, students in many countries face challenges posed by two shifts: the shift from guided to self-regulated learning and a shift within mathematics itself (Rach & Heinze, 2017). Focussing the second shift, mathematics at schools deals mostly with calculations based on real-world problems. Therefore, mathematics in school can be seen as a tool to solve such problems and technical aspects and procedures are in the foreground (Gueudet, 2008; Hoyles et al., 2001). A view of mathematics including proofs and formal definitions is partly covered in curricula but underrepresented in the classroom (Jordan et al., 2008; Witzke, 2015) At university, mathematics is treated as a scientific discipline (Engelbrecht, 2010; Gueudet, 2008). Accordingly, not the solution of specific application problems is of interest, but a theoretical perspective on mathematics as an axiomatic, deductive system. General concepts with abstract definitions and deductive proofs are central.

Conceptual and procedural knowledge

The aforementioned distinction between calculation and application versus axiomatic, deductive systems has been associated with the distinction between procedural and conceptual knowledge (Göller et al, 2022). Procedural knowledge is generally defined as “the ability to execute action sequences to solve problems”, whereas conceptual knowledge is defined as knowledge about the facts in a domain, their interrelations as well as the understanding of principles that govern a domain (Rittle-Johnson et al., 2001, p. 346). Therefore, conceptual knowledge is sometimes described as “knowing that”, whereas procedural knowledge is described as “knowing how” (Förtsch et al., 2018). While interrelated (see Schneider et al., 2011) both types of knowledge can be empirically separated and are associated with different uses. While procedural knowledge may be sufficient when solving familiar tasks, conceptual knowledge is often considered the basis for solving unfamiliar tasks. Moreover, procedural knowledge is restricted to specific types of tasks, while conceptual knowledge can be used flexibly in different types of situations and is generalizable (Rittle-Johnson et al., 2001). Besides generalizability, conceptual knowledge is further characterized by the possibility of representational change (Kaput, 1989), which is often a requirement for better understanding given mathematical problems and how procedural knowledge can be used to solve them. Regarding Linear Algebra, changing between a geometric view and an algebraic view can foster the understanding but is also an obstacle in itself (Gueudet-Chartier, 2004). Thomas et al. (2015) give an overview of students difficulties in Linear Algebra, especially an overreliance on an intuitive understanding (Waro et al., 2011) and a dominant view of basis as matrix manipulations (Thomas, 2011).

KUM-A level model for Analysis

Rach and Ufer (2020) described a level model of prior knowledge in Analysis for university courses. Based on a reanalysis of 21 items with 1553 students, four levels

were defined: the items on level 1 require procedural knowledge and knowledge of facts. They are formulated with little use of formal notations. To solve these items, students have to apply routine procedures. To solve items on level 2, students require conceptual knowledge to choose a solution procedure that is not obvious from the item text. Only routine changes of representations are included in level 2. Level 3 can be described by deeper conceptual knowledge, requiring integrated knowledge about different representations. To solve items on this level, students must link different representations or construct their own representations that are not given in the items. On level 4, formal notations and deductive reasoning are important. Empirical data underlined that level 3 was crucial for success in first-semester university Analysis courses (Rach & Ufer, 2020).

RESEARCH QUESTIONS

Aiming to create a model analogous to KUM-A for Linear Algebra, which can be used to better understand difficulties at the secondary-tertiary-transition and support students, this paper investigates two research questions:

- Is it possible to identify and describe a model of mathematical knowledge in Linear Algebra based on the KUM-A model from a theoretical perspective?
- Are the theoretically derived levels reflected by the empirical item difficulties from a scaling study and do students solution and difficulties when working on items of different levels correspond to their level classification?

METHODOLOGY

The Linear Algebra test consists of 25 multiple choice or open items. The items deal with the concepts of vector operations, orthogonal vectors, scalar products, linear combinations, linearly dependent vectors, straights, linear functions, linear equation systems, distances, and groups, which are taught in school as well as are important learning prerequisites for a Linear Algebra course in university (Halverscheid & Pustelnik, 2013).

To derive the level description for knowledge in Linear Algebra, we conducted several steps: first, we drafted a preliminary model for Linear Algebra based on the KUM-A level descriptions (Rach & Ufer, 2020). Second, we generated items for each level of this preliminary model. Third, preliminary cut-off scores for this model were determined based on the empirical difficulties of the items which were gained in an IRT scaling study with 182 students. This preliminary four-level model is described in Rach et al. (2021). Finally and for this contribution, we revised the preliminary four-level model based on the aforementioned theoretical considerations to improve its theoretical consistency. Modified level descriptions were formulated by the first author and revised by the second author. Afterwards, this new theoretical description was then again compared to the empirical item difficulties resulting from the scaling study to determine new cut-off parameters.

Written solutions of 34 students who worked on the test in their last school year were used to investigate the validity of our model, by considering whether students' solutions and difficulties corresponded to the theoretically derived level descriptions.

RESULTS

Table 1 shows the number of items assigned to each level of the revised model and the cut-off parameters for each level.

Table 1: Description for the four levels with number of items and cut-off parameters

Level	Description	Number of Items	Parameters
Level 1	Procedural knowledge, well known changes in representation, properties known from the real numbers	7	below −0.82
Level 2	conceptual knowledge, only one change in representation, concrete objects	8	−0.82 to 0.01
Level 3	Deepened conceptual knowledge, more than one change in representation, general objects	5	0.01 to 1.36
Level 4	Structural, conceptual knowledge, objects defined by their properties, working on mental representations	5	above 1.36

Comparing the items' allocations to the levels of the revised model with those of the preliminary model, 21 items were allocated to the same level in both models, while four items were shifted to another knowledge level by the revision. This leads to small shifts of the cut-off parameters compared to Rach et al. (2021).

Level 1 consists of items which require routine procedures that can be identified directly from the given problem or knowledge about facts, for example the definition of a circle or identifying the slope of a linear function. The procedures can be performed without conceptual understanding of the objects. Items require no changes of representations specific for Linear Algebra. The items do not contain coordinate representations. Switching between algebraic and geometric representations is only necessary in well-known situations. The items use little algebraic notation or can be translated into simple calculations by well-known routines. The correctness of general statements can be determined by calculations or properties known from the real numbers.

Items on *Level 2* require conceptual knowledge. For example, parameters have to be chosen such that given properties are fulfilled by some objects, such as distances and solution sets. In contrast to items on level 1, knowledge of facts or performing procedures is not sufficient. The items use some algebraic notations of concrete objects, for example to describe equations of lines in vector form or linear systems of equations. Changes between algebraic and graphical representations are limited to concrete objects, such as specific vectors or lines. In these cases, a coordinate system is given to perform such changes in representations. Changes in general situations,

given only by properties, are not necessary. Switching between geometric and algebraic representations of objects is also only dealt with in standard situations.

For items on *Level 3* deep conceptual knowledge that is not present as a fact is required. Different representations of objects and their properties are necessary to solve the items, sometimes more than one change of representations has to be performed. Items require self-chosen examples to evaluate statements of objects and their properties. Especially changes between geometric and algebraic representations are necessary to solve the items on this level. The geometric and algebraic properties have to be combined to solve the items. The representations of partially generic objects are changed but no operations on the objects need to be performed.

Items on *Level 4* require complex conceptual knowledge, which integrates algebraic and geometric representations of objects. Notations are on an algebraic level, for example the equation of a proportional function. Typically, more than one change of representations is necessary, for example from a verbal description to an algebraic description to solve the item, and afterwards to a geometric interpretation. The objects under consideration are not concrete objects but are given only by their properties. Therefore, representations of partially generic objects must be used. In addition, to solve the items, mental operations on geometric representations must be performed.

Discussion of example items

The following item (see Fig. 1) is an example for *level 2*: To solve the item, one has to use the properties of the scalar product. After identifying the correct procedure, the item can be solved by calculations with concrete vectors. Thus, one fact has to be combined with one procedure. The item can be solved by working with specific vectors given with specific numbers. Changing to a geometric representation (which is well known for given coordinates) of the situation can be helpful but not required.

Figure 1. Example item on level 2 (difficulty parameter –0.24).

Which pair of vectors $v_1, v_2 \in \mathbb{R}^3$ form a right angle?

☐ $v_1 = \begin{pmatrix}1\\1\\4\end{pmatrix}, v_2 = \begin{pmatrix}2\\5\\5\end{pmatrix}$, ☒ $v_1 = \begin{pmatrix}2\\-1\\5\end{pmatrix}, v_2 = \begin{pmatrix}2\\-1\\-1\end{pmatrix}$

☐ $v_1 = \begin{pmatrix}1\\-1\\2\end{pmatrix}, v_2 = \begin{pmatrix}1\\2\\-1\end{pmatrix}$, ☐ $v_1 = \begin{pmatrix}1\\0\\1\end{pmatrix}, v_2 = \begin{pmatrix}0\\1\\1\end{pmatrix}$

The difficulties found in students‘ answers to the item show that it requires more than a simple calculation since the connection between a right angle and the scalar product has to be made and may be missing (e.g. *“There is probably a Cosine-formula, which I always look up”)*. Some students also remembered false facts (e.g. *“The scalar product has to be 1 to have a right angle”)*.

As a second example (see Fig. 2), we discuss an item *of level 3*: deepened conceptual knowledge is necessary to solve the item by this group of students. Different concepts must be linked in a way that is typically not treated in school in this or a similar way. It is necessary to combine geometric properties, such as orthogonality, and algebraic properties, such as linear combinations and linear independence. The algebraic formulation of the item must be changed to a geometric representation to solve the item. The vectors are not specified, but only their properties are given.

Figure 2: Example item on level 3 (difficulty parameter 1.36, item is correct when all choices are correct).

Consider two vectors of $\mathbb{R}^2$, which are orthogonal to each other. Which of the following statements holds? Mark them.

	correct	not correct
- Each vector of $\mathbb{R}^2$ can be expressed as a linear combination of those two vectors.	☑	☐
- The zero vector can only be expressed as $0 \cdot v_1 + 0 \cdot v_2$ as linear combination of those two vectors.	☑	☐
- The scalar product of the two vectors is 0.	☑	☐
- The angles of intersection of v_1 with the x-axis and of v_2 with the x-axis are equally large.	☐	☑

One difficulty of the item is knowing the various concepts which are mentioned in this item. Mistakes show problems in translating the statements to a geometric representation, e.g. *"Just because two vectors are orthogonal, another vector cannot lay in the same direction"*. This statement indicates problems in mentally using a geometric representation, which points to problems in changing the representation. It also shows the problem of dealing with partially generic vectors which are only defined by their properties. This requirement characterizes the third level.

DISCUSSION

The proposed level model for prior knowledge in Linear Algebra as well as the adjoining empirical data highlight that it was possible to transfer the model for Analysis to describe knowledge in Linear Algebra. Mainly different depths of conceptual knowledge and the role of representational changes characterize the different levels of knowledge in Linear Algebra.

The examples of two items illustrate that the theoretical level description fit to difficulties of students solving the items. However, it would be desirable to further validate the model by analysing more written item solutions.

However, for the items in our Linear Algebra test, the representational changes are mostly between graphical and algebraic representations, which differs from the Analysis model. The high levels of demand in our test are less described by the number of integrated representations but by the usage of concrete numerical objects and generic objects given by their properties. A second limitation concerns our choice of items.

The model fits our 25 items but there are demands that are missing. For example, formal proof is missing in the items and is therefore not part of our model for Linear Algebra. However, solutions rates for the test item are between 4,7% and 96,0% covering a range of difficulties.

Our model allows for further investigation of study success in the first semester. Analogous to the approach of Rach & Ufer (2020), it would be interesting to see, which prior knowledge in Linear Algebra is related to success during the first semester at university and which level describes this knowledge. The description of demands regarding prior knowledge for the first semester of studies would also help the design of transition courses to support incoming students. In this context it is an open question to which extent the same levels of knowledge are achieved in Analysis and Linear Algebra by individual students.

REFERENCES

Engelbrecht, J. (2010). Adding structure to the transition process to advanced mathematical activity. *International Journal of Mathematical Education in Science and Technology*, 41(2), 143–154.

Förtsch, C., Sommerhoff, D., Fischer, F., Fischer, M., Girwidz, R., Obersteiner, A., et al. (2018). Systematizing professional knowledge of medical doctors and teachers: development of an interdisciplinary framework in the context of diagnostic competences. *Education Sciences, 8*(4).

Göller, R., Gildehaus, L., Liebendörfer, M., & Besser, M. (2022). Conceptual and procedural mathematical knowledge of beginning mathematics majors and preservice teachers. In C. Fernández-Verdú, S. Llinares, Á. Gutiérrez, & N. Planas Raig (2022), *Proceedings of the 45th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 307-314), PME.

Gueudet, G. (2008). Investigating the secondary-tertiary transition. *Educational Studies in Mathematics*, 67(3), 237–254.

Gueudet-Chartier, G. (2004). Should we teach linear algebra through geometry? *Linear Algebra and its Applications*, 379, 491–501.

Hailikari, T., Nevgi, A., & Komulainen, E. (2008). Academic self-beliefs and prior knowledge as predictors of student achievement in Mathematics: A structural model. *Educational psychology*, *28*(1), 59-71.

Halverscheid, S., & Pustelnik, K. (2013). Studying math at the university: Is dropout predictable? In A. M. Lindmeier & A. Heinze (Eds.), *Proceedings of the 37th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 417–424). PME.

Heublein, U. (2014). Student Drop-out from German Higher Education Institutions. *European Journal of Education, 49*(4), 497–513.

Hoyles, C., Newman, K., & Noss, R. (2001). Changing patterns of transition from school to university mathematics. *International Journal of Mathematical Education in Science and Technology*, 32(6), 829–845.

Jordan, A., Krauss, S., Löwen, K., Blum, W., Neubrand, M., Brunner, M., Kunter, M., & Baumert, J. (2008). Aufgaben im COACTIV-Projekt: Zeugnisse des kognitiven Aktivierungspotentials im deutschen Mathematikunterricht. *Journal für Mathematik-Didaktik*, 29(2), 83–107.

Kaput, J. J. (1989). Linking representations in the symbol systems of algebra. *Research Issues in the Learning and Teaching of Algebr*a, 4, 167–194.

Rach, S., & Heinze, A. (2017). The transition from school to University in Mathematics: Which influence do School-related variables have? *International Journal of Science and Mathematics Education*, 15(7), 1343–1363.

Rach, S., & Ufer, S. (2020). Which prior mathematical knowledge is necessary for study success in the university study entrance phase? Results on a new model of knowledge levels based on a reanalysis of data from existing studies. *International Journal of Research in Undergraduate Mathematics Education*, *6*(3), 375-403.

Rach, S., Sommerhoff, D., & Ufer, S. (2021). Technical Report – Knowledge for University Mathematics (KUM) and Mathematics Online Assessment System (MOAS).

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology*, *93*(2), 346-362.

Schneider, M., Rittle-Johnson, B., & Star, J. R. (2011). Relations among conceptual knowledge, procedural knowledge, and procedural flexibility in two samples differing in prior knowledge. *Developmental Psychology*, *47*(6), 1525.

Thomas, M. O., Freitas Druck, I. D., Huillet, D., Ju, M. K., Nardi, E., Rasmussen, C., & Xie, J. (2015). Key mathematical concepts in the transition from secondary school to university. In *The proceedings of the 12th international congress on mathematical education* (pp. 265-284). Springer, Cham.

Thomas, M. O. J., & Stewart, S. (2011). Eigenvalues and eigenvectors: Embodied, symbolic and formal thinking. *Mathematics Education Research Journal*, 23(3), 275–296.

Wawro, M., Sweeney, G., & Rabin, J. (2011a). Subspace in linear algebra: investigating students' concept images and interactions with the formal definition. *Educational Studies in Mathematics*, 78(1), 1–19.

Witzke, I. (2015). Different understandings of mathematics. An epistemological approach to bridge the gap between school and university mathematics. In *E. Barbin, U. T. Jankvist, & T. H. Kjeldsen (Eds.), ESU 7* (pp. 304–322). Danish School of Education.

PROSPECTIVE PRIMARY TEACHERS' UNDERSTANDING OF ONE-DIMENSIONAL PHENOMENA: LINE, RAY AND SECTION

Simone Reinhold and Bernd Wollring

Leipzig University & University of Kassel

Teachers' content knowledge is considered to play a meaningful role for teaching as (young) learners' development depends on their teachers' correct understanding and use of relevant terminology, especially in geometry. In the project we report partly in this paper, we intend to detect pre-service teachers' understanding of one-dimensional phenomena in Euclidean geometry, namely (straight) lines, rays and sections. Here, the prospective teachers were asked to spontaneously take notes concerning the different types of lines without any preceding instruction. The analyses reveal a great variety of individual notions, referring to everyday embedding or mathematical aspects and ranging from intuitive associations to formally adequate definitions. Aspects we assume to be beneficial as well as the impact of obstructive details are discussed.

INTRODUCTION

Considering that the concepts *line*, *ray* and *section* are basic topics in geometry and tackle fundamental ideas in various mathematical contexts, there has only been limited research concerning the understanding of these concepts, so far. Hence, there is a demand to deepen the research regarding (pre-service) teachers' and students' understanding of these concepts, comprising the interplay of their concepts during teaching and learning in the geometry classroom. Assuming that pre-service teachers' (PSTs') notions have a strong impact on their prospective students' learning, the initial study presented here intends to enlighten facets of PSTs' individual ground.

THEORETICAL FRAMEWORK

Subject matter knowledge: Issues in teacher education for geometry?

Building on the efforts of Shulman (1986), Ball et al. (2008) suggest that mathematical knowledge for teaching includes s*ubject matter knowledge* (SMK) and *pedagogical content knowledge* (PCK). Focusing on SMK, *horizon content knowledge* (HCK), *specialized content knowledge* (SCK) and *common content knowledge* (CCK) are defined as subdomains of SMK. HCK addresses knowledge concerning relations between topics, including familiarity with the curriculum. With a stronger focus on teacher's capability to analyze students' approaches and errors, SCK is considered to be specific for teaching. More generally, CCK comprises mathematical understanding and skills which are essential for teachers, but may also be important in settings *outside* the classroom (e. g. correct use of terms and notations). Aspects of CCK may not necessarily be mentioned in the classroom, but deficiencies in CCK can cause problems in SCK (Steele, 2013, p. 249). Therefore, even "unpacked mathematical knowledge" (CCK) which helps to distinguish between a term's mathematical meaning and its

2023. 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. 83-90). PME 46.

everyday reference is important "(…) because teaching involves making features of particular content visible to and learnable by students." (Ball et al., 2008, p. 400).

For more than 20 years, mathematics education research has revealed elements of PSTs' SMK, pointing at alarming lacks concerning content knowledge in geometry (Browning et al., 2014). As higher levels of SCK and CCK have been shown to correlate with stronger student outcomes in various studies, inadequacies of pre- and in- service teachers' SCK and CCK might be a potential source of students' difficulties in understanding (Steele, 2013). Hence, details and the development of PSTs' CCK obviously require ongoing and deep attention. For teacher education in geometry, this involves key facts and principles of Euclidean geometry, conceptual understanding of core concepts and the adequate use of terms (Jones, 2001; Sinclair et al., 2016).

Common content knowledge: Concepts in Geometry?

A concept involves the "(…) ideal representation of a class of objects, based on their common features" (Fischbein, 1993, p. 139). Representatives of a class of geometrical phenomena (one-dimensional figures, two-dimensional shapes or three-dimensional solids) can be visualized or grasped by visual (resp. haptic) perception (Franke & Reinhold, 2016). Thus, a concept includes a collection of characteristics which are considered to be relevant by an individual. A person may be able to verbally state these characteristics, but conceptual knowledge reaches beyond the capability of naming representatives or repeating formal definitions. It rather comprehends the perception, visualization and identification of distinctive properties which refers to individual mental images while thinking of a specific representative (Tall & Vinner, 1981).

Based on this, we have to consider that (prospective) teachers have encountered the concepts *line*, *ray* and *section* in various practical embeddings and formal contexts before they start teaching the topic themselves. They have met and used the terms in everyday contexts "over the years through experiences of all kinds" (Tall & Vinner, 1981, p. 152). In addition, they have most likely faced formal, precise and differentiated definitions, and have used this knowledge for geometric constructions. The theory of the *epistemological triangle* helps to understand this complex interplay

> (…) between the mathematical signs, the reference contexts, and the mediation between signs and reference contexts, which is influenced by the epistemological conditions of mathematical knowledge. (Steinbring, 2005, p. 22)

According to this notion, the concept of *line*, resp. *ray* and *section* and their abstract relations results from the individual interpretation of the related signs (e. g. one-dimensional representations in drawings) which have to be connected to an appropriate reference context or objects.

Line, ray and section as fundamental concepts in geometry

Basic activities with dynamic geometric environments (e. g. *GeoGebra* as one of the most popular DGE) practically illustrate that any geometric construction depends on the fundamental ideas of dots, lines, rays and sections and their interaction throughout

the digital design. Approaching from a historic and more theoretical perspective with Hilbert's *axioms of incidence* (*Axiome der Verknüpfung*; Hilbert, 1903, pp. 2-3), Hartshorne (2000) states:

> We simply postulate a set, whose elements are called *points*, together with certain subsets, which we call *lines*. We do not say what the points are, nor which subsets form lines, but we do require that these undefined notions obey certain axioms:
>
> **I1.** For any two points A, B, there exists a unique line l containing A, B.
>
> **I2.** Every line contains at least two points.
>
> **I3.** There exist three noncollinear points (that is, the points not all contained in a single line). (Hartshorne, 2000, p. 66)

Based on this, additional *axioms of betweenness* (*Axiome der Anordnung*; Hilbert, 1903, p. 4) underpin the definitions for the terms *ray* and *section*:

> If A and B are distinct points, we define the *line segment* $\overline{AB}$ to be the set constisting of the points A, B and all points lying between A and B. (Hartshorne, 2000, p. 74)
>
> Given two distinct points A, B, the *ray* $\overrightarrow{AB}$ is the set consisting of A, plus all points on the line AB that are on the same side of A as B. (Hartshorne, 2000, p. 77)

Freudenthal (1973) acknowledges this work, but most critically points at the limited use. Referring to the axiomatic approaches and the resulting formal definitions suggested by Hilbert and Pasch, he even states that they were complicated – useful for basic research but not for doing "geometry within them" (p. 402), and especially not beneficial for teaching geometry. Hence, suitable reference contexts (Steinbring, 2005) have to be considered: For example, the concept *ray* is fundamental for Euclidean plane isometries in geometry (e. g. shifting along a ray; rotation following the angle between two rays). Extending resp. shrinking a shape is visualized by the movement of certain points on rays. Grasping the aspect of infinity of a *line* may occur in contrasting this feature to the limits of a *section*. Furthermore, polygonal shapes consist of a closed path of sections. The concept *section* is also essential for the concept of measurement.

Reaching beyond: References of lines, rays and sections to arithmetic learning

There are various ways for visually representing numbers (Dehaene & Brannon, 2011). Yet, it is often suggested that the mental representation of numbers can take the form of a *(straight) line* in horizontal orientation. In primary, the external representation of natural numbers often refers to this *number line* (Obersteiner, 2018). The active use of this external representation plays a fundamental role for the development of mental representations. Hence, the concept of *rays* is used from grade one on – usually starting at zero on the left side, sometimes suggesting infinity by an arrow on the other side and putting marks on the line to represent the discrete numbers in the series of natural numbers. Later on, the concept of the *straight line* (with an open end on both sides) helps to represent whole numbers. Depending on the scale, each marking point on the number line defines the position of a number (in ordinal relation to other numbers) within this spatial structure. The difference between two natural numbers (resp. whole

numbers) is visualized by the distance of two points on this straight line – namely a *section* as part of this ray (resp. of the straight line).

RESEARCH QUESTIONS, DATA COLLECTION AND ANALYSIS

Aiming at contributions to an empirically grounded theoretical framework for PSTs' conceptual knowledge in one-dimensional geometry (with a focus on CCK), the main purpose of the partial study presented in this paper is to detect qualitative aspects of PSTs' understanding of *line*, *ray* and *section*:

- Which aspects are used by the PSTs in written notes to characterize their individual concepts of *line*, *ray* and *section*?
- (How) Are the different aspects used to indicate distinctions between the three concepts (e. g. with comments on the relationship or for an explicit distinction)?
- In how far does this lead to types of concepts?

The data collection for the analyses we share in this paper, focused on 135 written notes by PSTs in geometry courses at the end of their first year at a German university in 2022. During their second semester, the terms *line*, *ray* and *section* (German: *Gerade*, *Strahl*, *Strecke*) had not been addressed systematically, and no formal definitions had been negotiated. Yet, the terms might have been used when the PSTs talked informally about shapes and solids. With only general advice and without any further instruction, the PSTs were asked: "Please take a note: What do you understand by *line* (*ray*, *section*)?" Subsequently, the notes were negotiated in a communicative setting. Here, the PSTs presented their individual notes to peers and were able to revise the notes in the sense of collaboratively elaborated versions.

A research group, consisting of experienced and novice researchers (other PSTs shortly before final exams), developed a coding guideline mainly according to methods of Grounded Theory (Corbin & Strauss, 2015). Trying to detect qualitative aspects of the PSTs' individual understanding concerning *line*, *ray* and *sections*, the interpretation, coding and contrasting comparison of the data was supported by MAXQDA (2022). Furthermore, the research group passed phases of consensual validation (Beck & Maier, 1994) during their weekly interpretation sessions, resp. when using and elaborating the coding guideline for the analysis of these data sets.

For the analyses we share here, we focus on the initially taken *individual* notes and on the data we collected in this very special sample (PSTs in geometry classes with a geometry framing as a setting for the query). Yet, further and ongoing studies which are not reported here due to space limitations, widen our findings (e. g. data collection with primary students in grades 1 to 4, adults in non-academic settings, tasks including activities of drawing or sorting, using diagrams and digital technology). Initial inspiration for these and subsequent data collections was taken from earlier studies with secondary students with test items concerning the understanding of lines and

sections (Vollrath, 1998). For the analyses of this additional data, we took intercoder-reliability into account.

EXCERPTS FROM THE RESULTS

Codes and categories which emerged from the data turned out to be either more general or quite specific for *line*, *ray* or *section*. Beginning with the more specific findings concerning aspects the PSTs used in written notes to characterize their individual concepts, table 1 presents examples of aspects which were mentioned most frequently when PSTs offered notes for *section* (quotes throughout data translated by the authors).

Table 1: What is mentioned by PSTs when taking notes on *section*?

aspect	exemplary quotes from PSTs
beginning and end	has beginning and end (P37); line with defined beginning point and end (P39)
limitations	line limited by 2 points; limits on both sides (P13; P33)
measuring	distance from a to b; length between 2 points (P43; P33)

Although these codings seem to resemble, they sometimes occurred combined in the sense of distinct aspects which are obviously considered to offer complementary details (e. g. in the notes by P33). Yet, they did not necessarily occur coincidently, but were mainly found in seperate notes. Most frequently mentioned aspects for *line* which are displayed in table 2 spotted light on further aspects the PSTs referred to.

Table 2: What is mentioned by PSTs when taking notes on *line*?

aspect	exemplary quotes from PSTs
dimension and shape	line without curvature (P45); line without curves (P22); way without bending (P17); straight line (P16)
infinity	(runs) endlessly long (P23; P26), lies in space and can be endlessly long (P14); not limited by anything (P44)

When taking notes on *ray*, PSTs most frequently mentioned the *absence of aspects* which were stated for the other two concepts at the same time, namely the absence of an ending point (in contrast to *section*) and/or the absence of infinity (in contrast to *line*). In other words, identifying *limitations* and *infinity* seem to span two arbitrary aspects which are used deliberately in the notes which are underpinned by the *absence of aspects*. In addition, we found that the *absence of aspects* was also addressed to distinguish *line* from *section* (absence of beginning and end; e. g. P9). Yet, the absence of *infinity* was never mentioned to clarify the concept *section*. Most of these notes contain traces of Hilbert`s axioms of incidence and betweenness (see above), which PSTs obviously remind from their own textbooks in school. In addition and getting

back to our second question regarding the distinction between concepts, we can state that relationships are addressed either *explicitely* or via stating the absence of features (which are mentioned for one or two of the other concepts).

Furthermore, the aspect *dimension and shape* was never re-considered in our sample for the concept *ray* if it was mentioned for the other concepts at all, but *associations* occurred among notes for all concepts. For example, PSTs reminded themselves of the axes of the coordinate system or the time line and referred to the representation of the number line for the concept *ray*. The concept *section* was sometimes associated with a path which illustrates that the PSTs' not only included descriptions of a visualized static scenery, but used the visualization of movement (e. g. imagining a movement of themselves or imagining points "running" along a ray etc.)

P43: When I move from one point to another one, then the track I have covered is the section.

Regarding the question of different types of concepts, we found differences concerning the quality of the selected aspects and concerning the individual ways which were used to compose various aspects throughout the notes. Most frequently, the notes concerning the concepts line, ray and section could be characterized as:

- *single-feature-oriented*, with a tendency to add single aspects which were regarded as either complementary but often included redundancies,
- *genetic-oriented*, with a focus on constructive processes during the drawing of a one-dimensional representative,
- *relation-oriented in various entanglements*, mentioning shared aspects but distinguishing one concept from the other two concepts (or distinguishing one concept from only one of the other two) by specializing and/or contrasting (e. g. the absence of features, see above)

Yet, in some of the interrelating considerations we identified potential obstacles, as one or two of the concepts for *line*, *ray* or *section* might be adequate or at least promising whereas a related concept is not connected in an appropriate way:

P5: A section is the distance from one point to another one. A ray is a section which is directed from one point to a certain path.

P14: Ray illustrates that a line goes into a certain direction.

Surprisingly, neither the absence of *infinity* nor the quantitative aspect *shortest connection between two points* was mentioned for *section* by any of the PSTs in our sample. Instead, two quotes indicated severe misunderstandings or at least insecurity:

P17: A section has a starting point & no ending point (→ does not have to be straight?)

P20: A section describes the connection between two to several points. This [connection] does not need to be straight.

DISCUSSION, CONCLUSIONS AND OUTLOOK

Pursuing a qualitative empirical research paradigm, we constantly widened our perspective and shifted from stating the (non-)existence of formally correct and complete definitions to a closer look at the potential of individual wordings and the interpretation of references which we consider to be useful for teaching and learning in the classroom. Hence, we discovered similarities to the findings reported by Vollrath (1998) concerning grade 7/8 students' concepts of *line* and *section* (see tables 1 and 2). Connecting the concepts *line* and *section* with phenomena in real life (e. g. infinite lines in the universe, ruler, sections on a map; Vollrath, 1998, p. 25) also resembled our findings, although associations throughout our data were more likely restricted to mathematical contexts. Yet, according to our notion, these discoveries demand for an interpretation which reaches beyond (normative) expectations. We rather suggest to consider in detail how individual associations might be worth considering for the work in the classroom in the sense of adequate reference frames (Steinbring, 2005). Consequently and in a subsequent step, PSTs could and should discuss (resp. explore in diagnostic settings) to what extend these associations might be useful or obstructive for (young) students.

Furthermore, our analyses suggest that the concept *ray* plays an important role in the sense of interlinking the concepts *line* and *section*. This raises the question if addressing and exploring the relationship between all *three* concepts might provide a promising approach to foster PSTs' (and students') understanding. First analyses of data we gained from communicative settings with peer PSTs who negotiated and cooperatively revised their notes (see methods) point into that direction and suggest that communicative arrangements might help the individual PST to establish stronger reference contexts: Throughout negotiations, core contexts are identified with the aim to find colliding aspects or misunderstandings which then vanish in the PST-teams' revised notes. These enclose features of *relation-oriented* descriptions, which should be contrasted to *single-feature-oriented* and/or *genetic-oriented* descriptions as PSTs reflect and enrich their understanding, later on.

The study shows that acknowledging and appreciating aspects of individual concepts is meaningful and productive. We consciously limit this statement to the concepts *line*, *ray* and *segment*, but consider it quite possible that such an approach is also fruitful for other mathematical terms, especially for concepts in geometry.

REFERENCES

Ball, D. L., Thames, L., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? Journal of Teacher Education, 59(5), 389-407.

Beck, C., & Maier, H. (1994). Zu Methoden der Textinterpretation in der empirischen mathematikdidaktischen Forschung. In H. Maier & J. Voigt (Eds.), Verstehen und Verständigung (pp. 43-76). Aulis.

Browning, C. O., Edson, A. J., Kimani, P. M., & Aslan-Tutak, F. (2014). Mathematical Content Knowledge for Teaching Elementary Mathematics: A Focus on Geometry and Measurement. The Mathematics Enthusiast, 11(2), 333-384.

Corbin, J., & Strauss, A. (2015). Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. Sage.

Dehaene, S., & Brannon, E. (2011). Space, Time and Number in the Brain. Searching for the Foundations of Mathematical Thought. Elsevier.

Fischbein, E. (1993). The theory of figural concepts. Educational Studies in Mathematics, 24(2), 139-162.

Franke, M., & Reinhold, S. (2016). Didaktik der Geometrie in der Grundschule. Springer.

Freudenthal, H. (1973). Mathematics as an Educational Task. D. Reidel Publishing Company.

Hartshorne, R. (2000). Geometry: Euclid and Beyond. Springer.

Hilbert, D. (1903). Grundlagen der Geometrie (2nd edition). Teubner.

Jones, K. (2001). Teacher Knowledge and Professional Development in Geometry. In T. Rowland (Ed.), Proceedings of the British Society for Research into Learning Mathematics, 20(3), 109-114.

Obersteiner, A. (2018). Number sense across number domains: An integrated mathematics educational and cognitive psychological perspective. Habilitation Thesis TUM.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. Educational Researcher, 15(2), 4-14.

Sinclair, N., Bartollini Bussi, M. G., de Villiers, M., Jones, K., Kortenkamp, U., Leung, A., & Owens, K. (2016). Recent research on geometry education: an ICME-13 survey team report. ZDM, 48(5), 691-719.

Steele, M. D. (2013). Exploring the mathematical knowledge for teaching geometry and measurement through the design an use of rich assessment tasks. Journal of Mathematics Teacher Education 16(4), 245-268.

Steinbring, H. (2005). The Construction of New Mathematical Knowledge: An Epistemological Perspective. Mathematics Education Library, Volume. 38, Springer.

Tall, D., & Vinner, S. (1981). Concept Image and Concept Definition in Mathematics with Particular Reference to Limits and Continuity. Educational Studies in Mathematics, 12, 151-169.

Vollrath, H.-J. (1998). Zum Verständnis von Geraden und Strecken. Journal für Mathematikdidaktik, 19(2/3), 201-219.

THE ROLE OF LANGUAGE-AS-RESOURCE AND LANGUAGE-AS-POLITICAL IN COLLEGE MATHEMATICS COURSES

Jocelyn Rios

Colorado State University

This paper examines the role of language in shaping multilingual students' experiences navigating undergraduate mathematics courses. Drawing from what has been learned at the K-12 level, this paper uses the metaphors language-as-resource and language-as-political to capture the tensions that may exist in language diverse classrooms. Data stems from interviews with 28 multilingual students (whose home languages differ from the language of instruction), enrolled in introductory college mathematics courses. Findings demonstrate how students intentionally leveraged their home language as a resource in their mathematics courses, particularly as a resource for establishing community. At the same time, results highlight the role language played in students' marginalizing experiences in the classroom.

INTRODUCTION

College mathematics classrooms are becoming more linguistically diverse given the context of globalization and migration. In these spaces, students often bring different language resources that can be used to support their learning of mathematics. At the same time, languages carry different statuses in society (Civil, 2008), and students' language resources may not be equally valorized in the classroom (Planas & Civil, 2013). In the United States, discourses about English being the language of instruction may position native English speakers as the ideal language user in educational spaces (Subtirelu, 2015). Given this landscape, creating an equitable and inclusive classroom environment requires knowledge of how language can be leveraged to support college students' mathematical learning. It also requires recognition of the interconnectedness of language and status in the classroom.

The goal of this paper is to examine the role of language in undergraduate mathematics courses. Despite what has been learned at the K-12 level surrounding language and mathematics education (Barwell et al., 2016), little has been learned about supporting linguistic diversity in college math classrooms. Drawing from K-12 research, this paper explores if and how language serves as a resource for college students when navigating their mathematics courses. It also explores to what extent language shapes the political contexts of the classroom.

THEORETICAL FRAMEWORK

Planas and Civil (2013) present a framework for conceptualizing language in secondary mathematics classrooms using the metaphors language-as-resource and language-as-political. Language-as-resource represents the potential language has to facilitate multilingual students' learning of mathematics. Language-as-political

2023. 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. 91-106). PME 46.

represents the role that language plays in assigning status in the classroom, where some students are privileged are others are marginalized based on their language identity. Planas and Civil demonstrate this framework's relevance in understanding students' experience and the tensions that may emerge in multilingual classrooms. In the following sections, each metaphor is discussed and contextualized in the literature.

Language-as-resource

Language-as-resource is grounded in a sociocultural perspective, which emphasizes the role of cultural tools in the process of learning (Lerman, 2000). From this perspective, language is seen as more than just vocabulary and grammar. Rather, it is seen as a cultural tool for thinking, communicating, socializing, and making sense of the world (Moschkovich, 1999). Therefore, language-as-resource recognizes students' home languages as valuable tools for fostering mathematical learning. Studies in secondary mathematics education have examined how multilingual students use language, along with other resources like gesturing, to meaningfully engage in mathematical discussions during class (Moschkovich, 1999; Planas & Civil, 2013).

In addition, research has examined the ways that multilingual students use translanguaging to make sense of mathematics (Zahner & Moschkovich, 2011). Translanguaging captures how multilingual students often communicate by interweaving different linguistic resources (García, 2009). In the context of the math classroom, Planas (2021) defines translanguaging as "the creative use of language for mathematics teaching and learning as classroom participants make sense of their worlds and identities" (p. 10). Studies have shown that students use translanguaging in deliberate ways to support their cognitive and social engagement while working on mathematical tasks with peers (Planas, 2021; Zahner & Moschkovich, 2011).

Language-as-political

Language-as-political stems from a sociopolitical perspective on learning (Gutiérrez, 2013), emphasizing the role of language, power, and identity in shaping students' learning experiences. Planas & Civil (2013) highlight an example of language as political by discussing the idea of language valorization. This encapsulates how certain languages, like English, are positioned as being more valuable or useful than other languages. Native speakers of English are also often positioned as *superior* language users (Shuck, 2006). As such, linguistic features, like accent, can be used to signify outsider status. Furthermore, native speakers may feel less responsibile to communicate with non-native speakers when there are gaps in understanding during an interaction (Subtirelu, 2015). These political realities can shape how multilingual students negotiate identity and power in the classroom when interacting with others.

A growing body of research at the K-12 level has foreground a sociopolitical perspective on language. One of the main findings across the literature suggests that multilingual students tend to participate less during whole class discussions (Civil, 2008; Gorgorió & Planas, 2001; Planas & Civil, 2013). Planas & Civil (2013) document how bilingual students engaged in rich mathematical sense making using

their home language in a group setting. However, during whole class discussions, students felt the expectation that their contributions must always be in the language of instruction. This made them feel less comfortable participating. Students also felt pressure to speak in the language of instruction without mistakes. As a result, students' focus on grammar often interrupted their focus on mathematics. These examples demonstrate how language-as-political can impact on classroom participation.

Research Questions

The research questions that this study addresses are: *(1) In what ways did undergraduate students use their home languages as a resource for learning in their undergraduate mathematics courses? (2) In what ways did the political nature of language impact students' experiences learning mathematics?*

METHODS

The data analyzed in this paper stems from interviews with 28 multilingual students, whose home languages differed from the language of instruction (English). The languages represented in the sample included: Arabic (3), Bangla (3), Chinese (4), Farsi (1), Hindi (4), Korean (1), Spanish (9), Uzbek (1), and Vietnamese (2). Students either identified as being international students (18), immigrants (2), refugees (2), or transnational students (6), who lived near the border with Mexico and regularly moved between countries. In addition to their home language, students were selected for participation in this study based on their enrolment in either a pre-calculus or calculus course at the university where this study was conducted. The university was a large research university in the Southwestern, United States and a designated Hispanic-serving institution.

Semi-structured interviews were completed with each participant, lasting between one to two hours. Interviews focused on understanding students' experiences in their mathematics course and how language shaped these experiences. Students received $30 for completing an interview.

All interviews were transcribed in full and assigned pseudonyms. Transcripts were carefully read over and analyzed using a thematic analysis approach (Braun & Clark, 2006). This approach consisted of an initial round of content-driven coding, followed by a second round of data-driven coding. The content-driven coding focused on identifying instances in the data that exemplified language-as-resource and language-as-political. Definitions of these codes were developed based on the definitions in Planas and Civil (2013).

A second round of coding focused on identifying emergent themes among the examples of language-as-resource and language-as-political. An open coding system was used and descriptors for each code was developed. The final step of analysis focused on comparing, contrasting, and connecting codes, identifying themes across the data, and selecting quotes that were representative of each theme. The quotations included in this paper have been modified slightly for clarity. The main modification

consisted of removing partially repeated phrases. No modification impacted the meaning or feel of the quotes.

RESULTS

Language-as-resource

Students viewed language as a valuable resource that supported their learning and course engagement. By knowing more than one language, students felt that they could (1) be more flexible and creative when problem-solving, (2) have access to more resources like YouTube videos and textbooks (in both English and their home language), and (3) use language to connect with others in their course. This paper focuses on illustrating (3), as it was the most prevalent theme from the data. To do this, examples are shared from interviews with two students, Laura and Abhinav. These examples highlight the role of language in building meaningful connections with others and the positive impact that they had on students' course experiences.

Laura was enrolled in calculus I and was a local student, whose home language was Spanish. At first, Laura found it difficult to "click" with other students in her section and felt like they were not interested in connecting with her. Halfway through the semester, Laura started working on homework outside of class with a classmate, Javier. At one point, they worked on homework together over FaceTime since they were both at their homes. Laura recalled this experience, by sharing:

> I heard people in the background speaking Spanish, and then I heard a telenovela on, and I was like 'hold on a minute, you're telling me that you're Mexican?' and he was like 'yeah', and I was like 'whoa, this changes everything'.

Laura had wondered if Javier spoke Spanish, but "[hearing a telenovela] confirmed it". After this, Laura felt that her experience in calculus would be "a little bit better now".

Laura and Javier began regularly working on math homework together and helping each other learn: "I made a friend and I'm okay with the one friend that I have, because I know if I need help on a math problem, he'll help me". Laura and Javier's shared identity and background as Spanish speakers allowed them to develop "more of a connection, click" and support each other in valuable ways throughout the course.

Spanish not only facilitated their connection, but also served as a resource allowing them to communicate better. For example, Laura was able to utilize translanguaging when discussing mathematical ideas: "It's kind of second nature now. We just flip the switch and go from English to Spanish or Spanish to English. It's never just one thing, it's both". Laura did not use translanguaging because she was unable to sustain a conversation in one language. Instead, she described it as the natural way she communicates: it "flows better" and is more "comfortable".

In addition, being able to translanguage while working on mathematics allowed Javier and Laura to communicate more quickly, add humour, and "lighten up the atmosphere". This was particularly important to Laura, since doing mathematics often felt intimidating. Translanguaging helped created a space where Laura felt confident

and comfortable engaging with the course material, in a way that validated her linguistic resources.

For Abhinav, an international student from India, language also allowed him to connect with other Hindi speaking students in his calculus II course:

> Whenever I spot an Indian in the class, I feel a kind of connection with him. I immediately start thinking that he's my friend /…/ And whenever I talk to him, then I use Hindi /…/ So bilingualism has introduced that concept of familiarity.

In addition, language helped establish a close relationship between Abhinav and his calculus instructor, who spoke Hindi. Abhinav felt "more of a connection" with his instructor and saw him as a mentor he would "keep in touch with" after the semester ended. Abhinav regularly attended office hours, which was not a common practice among the other students interviewed in this study.

Moreover, communicating with his instructor felt easier, given their shared language background. For example, Abhinav did not feel "comfortable" speaking English in the classroom; however, speaking English with his professor felt "more comfortable":

> In front of a native speaker, I'm always conscious like, am I saying this correctly? Am I making a mistake here? /…/ But in front of him, I'm like, okay, he will understand what I'm gonna say. I don't need to be like super correct when speaking to him /…/ It makes it easier for me to communicate with him.

Abhinav was also able to speak Hindi during office hours. This created an affirming space where Abhinav could leverage his home language in his learning of mathematics. Laura and Abhinav's experiences highlight the role of language as a social resource, having the potential to help students build community that can support their learning and engagement in their mathematics courses.

Language-as-political

Student interviews also revealed ways that language was political in the undergraduate mathematics classroom and the tangible impact that that had on students' course experiences. Findings demonstrate (1) students' beliefs that only English should be used in the classroom, (2) that speaking English in a way that signifies non-native speaker status (i.e., speaking with an accent or making mistakes) would be viewed unfavourably in the classroom, and (3) that native speakers were less likely to meaningfully engage with the contributions of non-native speakers. This paper focuses on illustrating (3), by sharing examples from two students, Dep and Anita.

Dep was an international student from Vietnam taking calculus I. When Dep tried to participate during class, her peers and instructor often made her feel like they did not "have time" to "understand" what she was trying to say: "I tried to talking with them and like no one responds, so I stop talking". Dep recounted an experience where she asked her group: "anybody working on the question three, I have a problem with that." No one responded to her, which made her feel like her attempts at communication were

unintelligible. This also caused her to feel embarrassed: “I feel embarrassing because I’m thinking that maybe they don’t understand what I am saying”.

From talking with Dep during the interview, it was clear that she was an effective communicator in English. However, it seemed that her group members were not willing to put in effort to meaningfully engage with her contributions. Instead of listening and asking clarifying questions when needed, they ignored her contributions and made her feel like she was taking up time.

Dep also experienced similar behavior from her instructor:

> We [multilingual students] have different languages so there are some words we don’t [know]. We don’t know that much vocabulary compared to domestic students. So we are trying to find words to tell him. Sometimes I think he don’t have a lot of time talking with us. So that’s quite a problem.

The unwillingness of others to meaningfully engage in communication with Dep had tangible impacts on her participation. Being ignored by her group members made her decide to “stop talking” during class, even though she valued participating. Dep ultimately ended up dropping the course, which demonstrates how the politics of language can impact students’ academic trajectories in math or STEM.

Anita, an immigrant from Guatemala, had similar experiences working in groups in her calculus I course. Each week, Anita’s group had to take a timed group quiz, which required them to work together. The quizzes were heavily weighted. When Anita asked questions during the quiz, she was often positioned as “taking up time” and distracting the group from completing the quiz:

> Sometimes I wouldn’t know how to say a word and so sometimes my English would go like really bad, where they wouldn’t know what I was talking about. And they’re like ‘I’m sorry I don’t understand what you’re saying’. And they’re like ‘okay, well maybe forget your question and let’s just actually get to the problem’.
>
> I keep retrying to say what I meant, but he just gave up like ‘I’m sorry I don’t understand what you’re talking about’.

As these quotes evidence, Anita’s group members positioned her questions as *not relevant* to the problem they were working on. They also seemed to reject their role in making sure Anita’s contributions were understood (e.g., by asking questions, giving Anita space to rephrase her ideas, etc.).

Dep described being ignored by group members. In contrast, Anita’s group members were often rude and belittling when they did not understand what she was trying to convey. Anita shared that “one student was like ‘I’m sorry but we really need to get going’ but some others were like ‘dude, stop, we can’t do this’. And then that time I would feel like I’m so worthless”. In her calculus course, Anita felt like “an outcast” and woke up “heavier” on days she knew she had class. The marginalization that Anita experienced primarily stemmed from her peers’ unwillingness to meaningfully engage with her verbal contributions during group work. Like Dep, Anita also did not finish

the course and had to retake it over the summer, luckily with classmates who were kinder and more supportive.

DISCUSSION

The goal of this paper was to better understand the role of language in undergraduate mathematics courses. From interviews with multilingual college students, findings demonstrated the relevance of language-as-resource and language-as-political in post-secondary mathematics education contexts. On one hand, students clearly viewed their home language as a valuable resource, particularly for fostering social connections. Students leveraged language and their language identities to build community within their math courses. These communities became students' main sources of support in the course and the main way that they engaged with others. These communities also became spaces in their mathematics courses that affirmed their linguistic and cultural identities. By findings social resources that validated their home language, students also created spaces where learning and communicating about mathematics felt comfortable.

At the same time, students shared experiences in the classroom that demonstrated language-as-political. These examples often took place during group work. This is not surprising, since group work opens the classroom space for interactions and often makes identities more visible (Takeuchi et al., 2019). Findings exemplified how, during group work, native speakers did not always meaningfully engage with the contributions of non-native speakers. Students described instances where their contributions were ignored or belittled. Over time, these experiences inhibited their participation and sense of belonging in the course. This had tangible implications on students' experiences studying mathematics at the university.

More work is needed to explore how to better support language diversity in college mathematics classroom. Research can examine how to create more spaces where students can use their home language to learn mathematics (e.g., bilingual study sessions, recitation sections, study groups, etc.). These spaces could also have the potential to help students connect with peers and build community around a shared language identity. In addition, future research is needs which explores how to create classrooms norms that encourage equitable communicative responsibility. This describes the responsibility all speakers should assume to work together to establish mutual understanding during a social interaction (Subtirelu, 2015). Recognizing language as a resource, while also recognizing the sociopolitical context of the classroom, can help instructors be better equipped to build inclusive and equitable college mathematics classroom in this growing multilingual educational landscape.

REFERENCES

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101.

Civil, M. (2008). When the home language is different from the school language: Implications for equity in mathematics education [Paper presentation]. *The 11th International Congress of Mathematics Education (ICME)*, Monterrey, Mexico.

Barwell, R., Clarkson, P., Halai, A., Kazima, M., Moschkovich, J., Planas, N., Mamokgethi Setati, P., Valero, P., & Ubillús, M. V. (2016), *Mathematics Education and Language Diversity*. Springer Nature, Cham.

García, O. (2009). Multilingualism and translanguaging in the 21st-century. In T. Skutnabb-Kangas, R. Phillipson, A. K. Mohanty, & M. Panda (Eds.), *Social justice through multilingual education* (pp. 140–158). Multilingual Matters.

Gorgorió, N., & Planas, N. (2001). Teaching mathematics in multilingual classrooms. *Educational studies in mathematics, 47*(1), 7-33.

Gutiérrez, R. (2013). The sociopolitical turn in mathematics education. *Journal for Research in Mathematics Education, 44*(1), 37-68.

Lerman, S. (2000). The social turn in mathematics education research. In Boaler, J. (Ed.), *Multiple perspectives on mathematics teaching and learning* (pp. 19-44).

Moschkovich, J. (1999). Supporting the participation of English language learners in mathematical discussions. *For the learning of mathematics, 19*(1), 11-19.

Planas, N. (2021). Challenges and Opportunities from Translingual Research on Multilingual Mathematics Classrooms. In Essien, A., & Msimanga, A. (Eds.), *Multilingual Education Yearbook 2021: Policy and Practice in STEM Multilingual Contexts.* Springer.

Planas, N., & Civil, M. (2013). Language-as-resource and language-as-political: Tensions in the bilingual mathematics classroom. *Mathematics Education Research Journal, 25*(3), 361-378.

Shuck, G. (2006). Racializing the nonnative English speaker. *Journal of Language, Identity, and Education, 5*(4), 259-276.

Subtirelu, N. C. (2015). “She does have an accent but…”: Race and language ideology in students' evaluations of mathematics instructors on RateMyProfessors. com. Language in Society, 44(1), 35-62.

Takeuchi, M. A., Bryan, V., Valera, S., & Dadkhahfard, S. (2019). Immigrant Students’ Stories About Learning Mathematics in Groups. Canadian Journal of Science, Mathematics and Technology Education, 19(3), 237-253.

Zahner, W., & Moschkovich, J. (2011). Bilingual students using two languages during peer mathematics discussions:¿ Qué significa? Estudiantes bilingües usando dos idiomas en sus discusiones matemáticas: What does it mean. In Telléz, K., Moschkovich, J., Civil, M. (Eds.), *Latinos/as and mathematics education: Research on learning and teaching in classrooms and communities* (pp. 37-62). Information Age Publishing.

EMBODIED CURIOSITY: A FRAMEWORK FOR MATHEMATICAL MEANING-MAKING

Sheree Rodney

Ontario Tech University

This research investigates the role of curiosity and Geometer's Sketchpad in developing mathematical meanings with grade nine students (14 -15years old). I introduce Embodied Curiosity, a theoretical framework that speaks to the connectivity between trait-curiosity and digital technology, and use it to exemplify the attributes of learning that can be interpreted as observable curiosity while providing insights about learners' mathematical meaning-making. The research findings show the body links curiosity to technology in ways that provide opportunities for mathematical meanings to be constructed.

Keywords: curiosity, digital technology, embodiment, mathematical meaning

INTRODUCTION

There is high research interest in how to encourage students' meaning-making using technology (Schnaider & Gu 2022; Manshadi, 2021) and more so on how to foster students' meaning-making as teachers in mathematics (Brodie, 2010; Jaworski, 2015). The basis of this research is to illustrate how we can gain insights into learners' meaning-making processes when the focus is on embodied learning. This study introduces a new theoretical framework which I called Embodied Curiosity and uses data from one episode (extracted from a wider data pool) to validate the framework as a potential theoretical consideration for mathematics learning. Additionally, I show how the Embodied Curiosity framework is conceptualized using descriptions of its main principles, and visualization of the interconnectedness among the fundamental elements on which the framework is grounded.

In this study, curiosity, or trait-curiosity, is reconceptualized not solely as an intrinsic motivation but rather as a relationship with the external. I have adopted Loewenstein's (1994) curiosity definition that curiosity is "as a form of cognitively induced deprivation that arises from a gap in knowledge or understanding" (p. 75). Implying that curiosity is a realization of the imbalance between what we know and what we want to know. He draws attention to how problematic it is to theorize about curiosity and argues that previous definitions did not capture certain significant characteristics of it. One such example is its intensity (the pain of not having information) and another, is its transience (how short-lived curiosity is). Furthermore, I take mathematical meanings to be the product of students' beliefs, use of symbols, verbalisation, visual representations, and making connections to mathematical concepts while doing mathematics. The Geometer's Sketchpad was specifically chosen for this research because of the dynamic visualization and affordances it provides. The research addresses the following research questions: To what extent does the body triggers

2023. 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. 99-106). PME 46.

curiosity? And what insights can be gained about learners' meaning-making in mathematics through engagements with Sketchpad?

CONTEXT AND BACKGROUND

Curiosity is immeasurable and cannot be seen with the naked eye. For this reason, it has been problematic to define. In many researches, curiosity is seen as an internal motive that influences human behaviour and fosters active learning (Oudeyer, Gottlieb & Lopes, 2016; Kidd & Hayden, 2015). However, Berlyne (1954) suggests two types of curiosities, one that increases the perception of a stimulus (perceptual or sensation-seeking) and the other that reinforces knowledge (epistemic or knowledge-seeking). He argues that these two types of curiosities inform not only how we become curious, but also, why we are curious. Berlyne's dualistic (epistemic vs perceptual) approach is meaningful to this research because it implies that curiosity is motivated internally and externally which sets up an intersectional space for curiosity and digital technology–a space that is of interest to this research.

Sketchpad

Sketchpad is dynamic geometry software (DGS) that combines technology and mathematics in teaching geometry. The software has been used to create, explore and analyze a wide range of mathematical concepts in algebra, geometry, and calculus among other mathematical areas (Bakar, Tarmizi, Ayub & Yunus, 2009). It includes Euclidean geometry tools such as point, circle, polygonal shapes, lines, line segment, and line ray, all of which are used to produce geometric drawings and constructions. *Sketchpad* is designed to encourage students to discover how geometrical shapes are related, which contributes to the development of their own mathematical meanings. Due to the dynamic nature of the software, learners experience hands-on manipulation of mathematical concepts while identifying patterns, making conjectures, and producing proofs at a faster rate than they would in static environments. Furthermore, the draggability feature of the software allows the learner to examine examples and non-examples of geometric constructions on the same screen at the same time. In research conducted by Arzarello et. al's (2002) who analyzed learner's dragging practices, found that learners engage in certain dragging modalities when they perform a geometric tasks using DGS: Wandering dragging (aimlessly moving the points on the screen), Bound dragging (moving a point that is linked to an object), Guided dragging (dragging the basic points to give it a particular shape), Dummy Locus Dragging (dragging a point to maintain a previously discovered property), Line Dragging (drawing new points to keep the shape of the figure), Linked Dragging (connecting a point to an object and attaching it to the object), and Dragging Test (moving draggable points to test if the shape main its properties).

Bearing these dragging practices in mind, I have extended Arzarello et. al's modalities to include **curious dragging**. Curious dragging includes these modalities with something extra– temporality, speed, and emotion. For example, Arzarello et. al's bound dragging occurs when the learner moves a constructed point that is linked to an

object. If this is done in a slow meticulous manner or fast with excitement it can be interpreted as curious dragging. I demonstrate this further in the analysis of the episode.

Mathematical-Meaning Making

According to Radford et. al (2011), the learning of mathematical concepts is consistently changing which creates space for the development of creative and critical involvement in the way children learn. This shifts the focus away from traditionally recognizing the teacher as the sole dispenser of knowledge, to embrace the fact that learners are capable of constructing their own knowledge. Based on this, there is a constant need to better understand how meanings are constructed in mathematics. To this end, the mathematical meaning-making process involves the way learners think and communicate about mathematical concepts when they interact in a social context. Radford et. al suggest that this can be "accomplish[ed] through, written, oral, bodily, and other signs" (p.150). Since this research is concerned with the understanding of geometric terms relating to the circle theorem, the mathematical meaning-making process emphasizes how learners come to understand the connections between concepts pertaining to the circle. I argue that these meanings emerged both explicitly and implicitly providing implications for pedagogical and research interests.

THEORETICAL FRAMEWORK

Researchers have proposed that the body is more than skin-bound and plays an important role in learning mathematics. For example, Gol Tabaghi & Sinclair (2013) suggest that learning takes place when speech, body movement, gestures, and material work together in a harmonious relationship. Similarly, Embodied Curiosity, the theoretical framework that informs this study, focuses on a network of four fundamental elements (curiosity, body movement, digital technology, and mathematical meanings). The framework is taken as a stratigraphic structure (Figure 1) with semi-permeable layers similar to that of a biological cell. The layer of curiosity is considered the main component and starting point of the structure and is usually triggered by students' wondering or uncertainty. The semi-permeable nature of the layers allows for the distinction of examples and non-examples of instances or factors that connects each layer to the other.

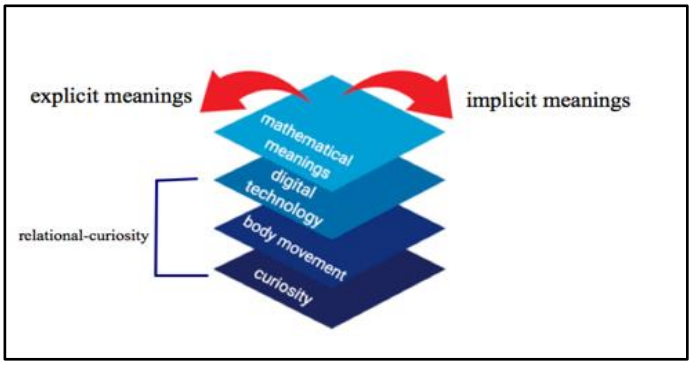


Figure 1: The Embodied Curiosity Model

For example, while visual fixation may be considered an emerging curiosity, its movement throughout the framework could be restricted because visual fixation is also associated with other emotions like interest and attentiveness. Likewise, a learner sitting in an upward position while clicking the mouse cannot be considered emerging curiosity unless this movement is followed by a lean forward while attending to something specific with the technology. This implies that in order for curiosity to be

triggered, developed, and sustained there must be a combination of factors and free movement of certain factors from one layer to the other.

In this model, the learner demonstrates curiosity when they ask questions that imply a gap in their knowledge or express uncertainty about a concept. When this occurs the learner usually seeks to fill that gap through exploration with the technology. In doing so, the body reacts in certain ways creating an interweaving connection among curiosity, body movement, and technology. I call this connection **relational-curiosity,** which is the first principle of the Embodied Curiosity framework. It is important to note that this type of curiosity is not controlled by the learner, but instead, relies heavily on both the learner and technology working hand-in-hand. Due to the displacement of learners from the central position within an interaction, the second principle of the Embodied Curiosity framework is that it is **temporal and emergent**. This means Embodied Curiosity is unplanned and emerges in real-time, while the third principle has to do with Sketchpad's dragging capability. I suggest that learners engage in a dragging modality called curious dragging. This type of modality is an extension of Arzarello et. al dragging modalities and it also involves time, speed, and emotions. Based on the model, when a learner experiences relational-curiosity two types of mathematical meanings emerge; those that are explicit (verbalized by the learners) and implicit, which are not clearly expressed. Explicit meanings symbolize that mathematical meanings are constructed while the implicit meanings can be used by the teacher to resuscitate the Embodied Curiosity process.

METHODOLOGY

Video recordings of classroom observations and students' work from a high school located in the Caribbean form part of the data used in this study. The school is referred to as School Y for anonymity and was purposely selected because of its technology-rich environment– a fully equipped computer lab with at least 30 computers. The participants (pseudonyms: Kyle, Jeff, and Ali) were randomly selected from a class of grade nine students (ages 14 and 15 years old). At the time of the research, the students were beginning the geometry section of the mathematics curriculum and the topic circle geometry theorem was chosen to accommodate the research. Furthermore, the students had prior exposure to geometry in previous grades and were already familiar with the basic properties of the circle, albeit in a static environment. However, they were doing circle geometry theorem for the first time, and most significantly, both the teacher and the students were using *Sketchpad* for the first time as well.

The classroom observations were done over a three-week period, where students engaged in tasks using only *Sketchpad*. They were allowed to collaborate with each other, either working on individual computers or by sharing one computer in pairs or triads. The duration of the classroom sessions was forty-five minutes and the classroom interactions were video recorded as well as students' work was saved and retrieved from the schools' server after each session. The mathematics curriculum is offered approximately five times per week. However, I was accommodated for two sessions

(90 minutes) per week in the school. During the non-research sessions, the teacher engages students in simple geometry tasks, such as constructing and classifying triangles based on their properties. In this way, both teachers and students became more proficient in using *Sketchpad* and the students were able to reinforce the knowledge they may have experienced in previous lessons. Data from video-recorded classroom sessions and students' work were analyzed in a continuous, iterative manner to ascertain how and when curiosity was triggered after which, episodes were interpreted for evidence of relational-curiosity (students' uncertainty/ wondering-body movement-Sketchpad). Finally, these selected episodes were analyzed to locate scenarios when mathematical meanings emerged and the episode presented in this research is one of eight examples that were identified.

RESULTS AND ANALYSIS

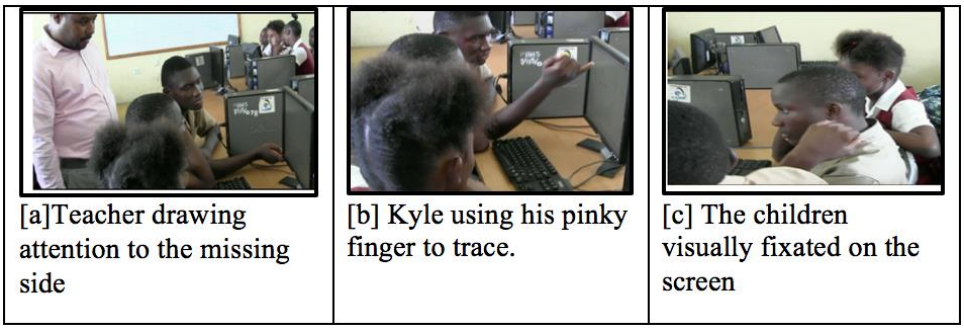
The data were analyzed with an eye on how curiosity could be identified and leveraged for mathematical meaning-making, as well as, the role of *Sketchpad* and its affordances in the process. The data was examined through the stratigraphic lens of the Embodied Curiosity process with consideration on the three main principles. However, the focus of the analysis in this study is on the explicit meanings and not the implicit ones. The episode represented in this research was selected from six similar examples.

Episode one: The smallest finger holds the biggest secret.

Kyle, Jeff and Eli, are working together on a task using the shared-computer arrangement. They had been given the task to identify the type of triangle formed when two random points on the circumference are connected to the centre point of the circle. This is their fifth session and their interaction with *Sketchpad* has improved significantly. In performing this task, Jeff, performs the construction in the following order; circle, two random points on the circumference and then two line segments connecting centre point and points on the circumference. This sequential action produces an incomplete triangle inscribed in the circle shown in Figure 2 [a]. The children seem satisfied that they constructed a triangle in the circle.

Figure 2: [a] Incomplete triangle [b] Completed triangle

After a few minutes had pass, Kyle, who is sitting left of the computer (Figure 3), invites his teacher to their computer station and asks, "Sir, something like this?" Implying that Kyle is uncertain about the construction. His teacher notices that the triangle is incomplete (only sides AC and BC were drawn) and said, "But – but it needs one more <pause> that – that's a triangle now?" The teacher's pause mid-sentence is an indication that he refrains from telling them exactly what was missing.



[a]Teacher drawing attention to the missing side	[b] Kyle using his pinky finger to trace.	[c] The children visually fixated on the screen

Figure 3: Students exploring the isosceles triangle

Kyle then leans forward (Figure 3[a]) to investigate the missing information while Jeff (centre) and Eli (right) are visually fixated on the screen. In keeping with the Embodied Curiosity framework Kyles' uncertainty and leaning forward action implies that curiosity is triggered and the process of **relational-curiosity** is activated. The missing side of the triangle signifies the gap in their knowledge. However, Jeff and Eli's visual fixation is restricted because there is no telling if they are curious. As such, Kyle was able to progress from one layer of the model to the other but the others did not. It would appear that the students conceive arc AB as one side of the triangle. In an attempt to fill this gap, Kyle used his pinky finger to perform a tracing action on the screen (Figure 3 [b]). This action shows the connection between the body and digital technology, but most significantly, the development of the idea that the third side of the triangle should be a line segment rather than a curved arc as it appears in Figure 2 [a]. This action also helps Jeff and Eli to complete the diagram seen in Figure 2[b]. With the completed triangle on the screen, Kyle beckoned to his teacher a second time to join them at their computer station. This time, after looking at the construction the teacher asked:

1.Teacher: "So how will you know what type of triangle is that? What do you notice?

2. Eli: The sides are equal (hovering the cursor over the two radii)

3.Teacher: hmmm~~ and when two sides are equal what kind of triangle is that?

4. Kyle: <contemplating out load> equilateral ~~scalene (while using his pinky finger to trace several triangles on the top of his desk) What is the name of this triangle again?

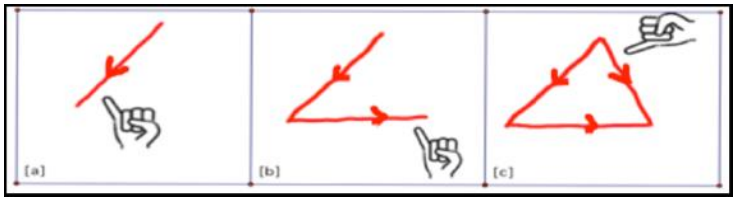


Figure 4: Kyle's finger motion drawing the scalene triangle on the desk top

The teacher's leading question, "What type of triangle is that?" (Turn 1), and the probing question "what do you notice?" (Turn 1), helped Kyle to re-examine his previous knowledge. Again, he used his pinky finger (Figure 4), this time, to draw several triangles on the desk while he decides on which triangle is on the screen (Turn 4). There seemed to be an internal conflict with associating the vocabulary word with the property of the triangle. Again, this is an indication that there is a knowledge gap

and that Kyle is uncertain about which triangle is on the screen. On the surface level, Kyle's knowledge seems to shift focus from the necessary; knowledge he can arrive at from previous knowledge (the two sides having the same length), to the arbitrary; knowledge he needs help with (the vocabulary word) as explained by Hewitt (1999). After a few minutes had passed, Eli decides to measure the attributes of the triangle "to find out if the sides are equal" and test her conjecture in Turn 2. Identifying that this was so, Kyle quickly dragged point A along the circumference, making point B invariant and performed Arzarello et. al's dummy locus dragging, that is, dragging a point to maintain a previously discovered property. Upon doing so, they recognise that the side lengths and two angles maintain equality throughout (Figure 2 [b]). In addition, Kyle engaged in **curious dragging** as his dragging motion is swift, and with excitement as the magnitude of the side lengths and angular measures changes. However, when point A coincides with point B on the circumference, Kyle is meticulous and cautious in his action as if he is concerned about disrupting the properties and conjectures he already discovered. The students are able to say the sides and angles were equal but unable to attach the name of the triangle to the properties.

DISCUSSION

The purpose of this study was to introduce a new way of thinking about how curiosity can be leveraged for the development of mathematical meanings and to give an account for the mediatory role digital technology plays in the process. The results showed that students wondering and uncertainties manifest through questions and body movements which are possible physical markers for observable curiosity. It was also evident that **relational-curiosity**, a construct of the Embodied Curiosity framework, unfolded when the students performed tasks relating to the circle geometry theorem. Additionally, the analysis showed that learners were able to verbalize the property (two equal sides and two equal angles) of the isosceles triangle explicitly but were unable to match the vocabulary word to the meaning. Kyle's use of his pinky finger on the computer screen showed how the technology prompts the body to react in real-time which implies the **temporal and emergent** nature of the Embodied Curiosity process. Furthermore, this action guaranteed communication about the mathematical meaning was inclusive perhaps for all the students and filled the information gap (the third side of the triangle) that created the imbalance in the first place. Another interpretation could be that Kyle reified the missing third side as something concrete and, it seemed his pinky finger was entrapped with the technology and it was in this entrapment that caused his curiosity to be satisfied. Once relational-curiosity was present the opportunity appeared for mathematical meanings to be developed. This was evident when Kyle, through the draggability of *Sketchpad*, engaged in **curious dragging** which allowed them to match the changes in side length and angular measures with their visualization of the triangle on the screen. Despite, their challenges in associating this property with that of an isosceles triangle, they were able to recognize that the third side was a line segment rather than an arc and that this special triangle has two equal sides and two equal angles. They could also tell that by changing the position of a point

on the circumference changes the magnitude of the side length and angular measure but does not disrupt the property of the triangle.

REFERENCES

Arzarello, F., Olivero, F., Paola, D., & Robutti. (2002). A cognitive analysis of dragging practices in cabri environment. *International Journal on Mathematics Education, 34*(3), 66–72.

Bakar, K.A., Tarmizi, R. A., Ayub, AFM., & Yunus, ASM. (2009). Effect of utilizing geometer's sketchpad on performance and mathematical thinking of secondary mathematics learners: An initial exploration. *International Journal of education and information Technologies, 3*(1), 20–27.

Berlyne, D. (1954). A theory of human curiosity. *British Journal of Psychology, 45*(3), 180–191.

Brodie, Karin. (2010). Pressing dilemmas: Meaning-making and justification in mathematics teaching. *Journal of Curriculum Studies*. 42. 27-50.

Gol Tabaghi, S. & Sinclair, N. (2013). Using dynamic geometry software to explore eigenvectors: The emergence of dynamic–synthetic–geometric thinking. *Technology, Knowledge and Learning, 18*(3), 149–164.

Hewitt, D., (1999). Arbitrary and necessary part 1: A way of viewing the mathematics curriculum. For the learning of Mathematics. 19(3), 2-9.

Jaworski, B (2015). Mathematics meaning-making and its relation to design of teaching. PNA, 9(4), 261-272.

Kidd, C., & Hayden, B. (2015). The psychology and neuroscience of curiosity. *Neuron, 88*(3), 449–460.

Lakoff, G. & Núñez, R. (1997/2000). *Where mathematics comes from: How the embodied mind brings mathematics into being*. New York, NY: Basic Books.

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin, 116*(1), 75–98.

Oudeyer, P.-Y., Gottlieb, J., & Lopes, M. (2016). Intrinsic motivation, curiosity and learning: Theory and applications in education technologies. *Progress in Brain Research, 299*, 257–283.

Manshadi, S. (2021). Meaning making in a sixth-grade mathematics classroom through touch screen technology. *International Journal of Mathematical Education in Science and Technology:* Taylor & Francis, 1-17.

Radford, L., Edwards, L., & Arzarello, F. (2009). Beyond words. *Educational Studies in Mathematics, 70*(2), 91–95.

Scher, D. (2000). Lifting the curtain: The evolution of the Geometer's Sketchpad. *The Mathematics Educator*, 10(1), 42–48.

Schnaider, K., & Gu, L. (2022). Potentials and challenges in students' meaning-making via sign systems. Multimodal technologies and interaction.6(9), 2-22.

HIGH SCHOOL STUDENTS' PERCEPTIONS OF THE RELEVANCE OF MATHEMATICS IN HIGHER EDUCATION

Dunja Rohenroth, Irene Neumann and Aiso Heinze

IPN - Leibniz Institute for Science and Mathematics Education, Kiel, Germany

Mathematics plays a central role in the STEM subjects (science, technology, engineering and mathematics) and also in a large number of other subjects. However, many students start their university studies with inadequate expectations regarding the needed mathematics. The present paper reports on a study investigating N=984 German high school students' perceptions of the relevance of mathematics for university studies by contrasting the subjects economics, education, medicine, and physics. Both, students' perception of the relevance of mathematics in general as well as of specific required mathematical aspects were assessed for the different subjects. The results indicate that many high school students underestimate the relevance of mathematics, especially in education studies.

INTRODUCTION AND THEORETICAL BACKROUND

Decades ago, Niss (1994) attested mathematics the so-called relevance paradox „formed by the simultaneous objective relevance and subjective irrelevance of mathematics" (p. 371). Previous studies have shown that already (high) school students perceive a subjective irrelevance of mathematics (e. g. Brown et al., 2008; Onion, 2004). For example, Brown et al. (2008) report 16 years old high school students say "the amount of insignificant maths work that I will NEVER use is quite big" (p. 11) or to ask "who needs to know trigonometry in everyday life?" (p. 11).

This unseen relevance is in contrast to the fact that a wide range of subjects to be studied at universities includes mathematics, certainly in the field of STEM (e.g., Deeken, 2020) but also beyond (e.g. in economics, medicine or psychology, Mulhern & Wylie, 2016; Weintraub, 2002). So, it does not come as a surprise that about 70 % of US bachelor students earn credits in mathematics courses (Douglas & Salzman, 2020) and that in Germany more than 80 % of the freshmen are expected to bring mathematical knowledge and abilities from school (Neumann et al., 2021). Two recent Delphi studies identified which mathematical prerequisites are expected from incoming students, from the perspective of university instructors (Deeken et al., 2020; Rohenroth et al., under review). Expected aspects span across mathematical content, mathematical processes, views about mathematics as a scientific discipline, and personal mathematics-related characteristics. Although these studies show that university instructors have – sometimes even quite high – expectations, many students take up a study holding inadequate expectations with respect to the required mathematics (e. g., Venezia et al., 2003). Missing information and inadequate expectations about required mathematics, in turn, may likely result in frustration during

2023. 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. 107-114). PME 46.

the studies and, at the worst, in student drop-out. In fact, many university students report that they have been surprised about how much mathematics and which particular mathematical knowledge and abilities are required in their studies (e. g., Darlington & Bowyer, 2016).

RESEARCH QUESTIONS

The present study aims to complement these retrospective studies by contributing a prospective perspective, that is, by investigating high school students' views about mathematics for their future university studies. In particular, we address the following research questions: (1) How do high school students perceive the relevance of mathematics for different subjects to be studied? (2) To what extent are high school students' views about the required mathematical aspects adequate given university instructors' expectations? (3) To what extent does the adequateness of students' views relate to the relevance of mathematics they see for particular subjects of study?

METHOD

The sample comprises N = 984 high school students from upper secondary level (533 female, 434 male, 13 diverse, 4 no response) from four German federal states who voluntarily completed an online questionnaire sent to their teachers or school principals. First, the students were asked to rate the importance of mathematics for the university studies of economics, education, medicine, and physics (1: "no importance" to 10: "very high importance"). Then, participants were presented specific mathematical prerequisites, which had been identified in the previous Delphi studies MaLeMINT and MaLeMINT-E (Deeken et al., 2020; Rohenroth et al., under review). These mathematical prerequisites addressed mathematical content (basic aspects from lower secondary as well as aspects from calculus, vectors and matrices, and stochastics) and mathematical processes (e. g., mathematical modelling). For each aspect and for each of the four study subjects, students were asked to indicate if they expect that the aspect is required for a study of the respective subject (yes/no). To not overburden students, they were assigned one of two overlapping lists of 24 aspects.

For data analysis, we compared high school students' ratings with expectations previously indicated by the university instructors. For each student, a mathematical learning prerequisite was scored to be *adequate* if the student indicated the aspect as (not) necessary and the aspect had been identified as (not) necessary by university instructors in the respective subject as well. If a students' rating differed from the university instructors' expectation, the prerequisite was scored as *underestimated* (students: not necessary; instructors: necessary) or *overestimated* (students: necessary; instructors: not necessary). Then, we determined the percentage of prerequisites scored as adequate, underestimated and overestimated for each high school student and each of the four subjects. Given that the percentages of overestimated prerequisites were very low (0 % - 4,1 %), we focused on the percentage of adequately scored prerequisites for further analyses.

RESULTS

RQ 1: Perceived relevance of mathematics for different subjects of study

Investigating the perceived relevance of mathematics in general, we found high school students to clearly differentiate between the four subjects of study (Economics: $M = 7.5$, $SD = 1.9$; Education: $M = 3.5$, $SD = 1.9$; Medicine: $M = 5.9$, $SD = 2.1$; Physics: $M = 9.4$, $SD = 1.3$, see also Table 1). As expected, students rated the relevance of mathematics highest for the study of physics, followed by the studies of economics and medicine. The relevance of mathematics for the education studies was viewed quite low. In fact, about the same percentage of students (i.e., more than 85 %) indicated that mathematics has no or a (rather) low relevance for the studies of education as they viewed mathematics (very) highly relevant for physics (Table 1).

	No or low relevance (1-2)	Rather low relevance (3-5)	Rather high relevance (6-8)	(Very) High relevance (9-10)	
	abs. (in %)	abs. (in %)	abs. (in %)	abs. (in %)	Total
Economics	13 (1,4 %)	137 (14,7 %)	480 (51,4 %)	303 (32,5 %)	933
Education	323 (34,7 %)	465 (49,9 %)	130 (13,9 %)	14 (1,5 %)	932
Medicine	48 (5,1 %)	350 (37,5 %)	437 (46,8 %)	98 (10,5 %)	933
Physics	10 (1,1 %)	11 (1,2 %)	108 (11,6 %)	806 (86,2 %)	935

Table 1: Perceived relevance of mathematics for different subjects.

RQ 2: Adequateness of students' views about the required mathematical aspects

Figure 1 displays the distribution of students with respect to their percentages of adequately scored learning prerequisites. It shows that the above findings regarding the relevance of mathematics in general also hold true for the specific learning prerequisites. More than 75 % of the students adequately scored at least 70 % of the prerequisites as necessary for the study of physics. In contrast, the vast amount of high school students underestimated most of the prerequisites required for the study of education. Such obvious trends are missing for the studies of economics and medicine. However, it has to be noted that only about a third (economics) or even less (medicine) of the students adequately scored at least 70 % of the prerequisites required for the respective subject.

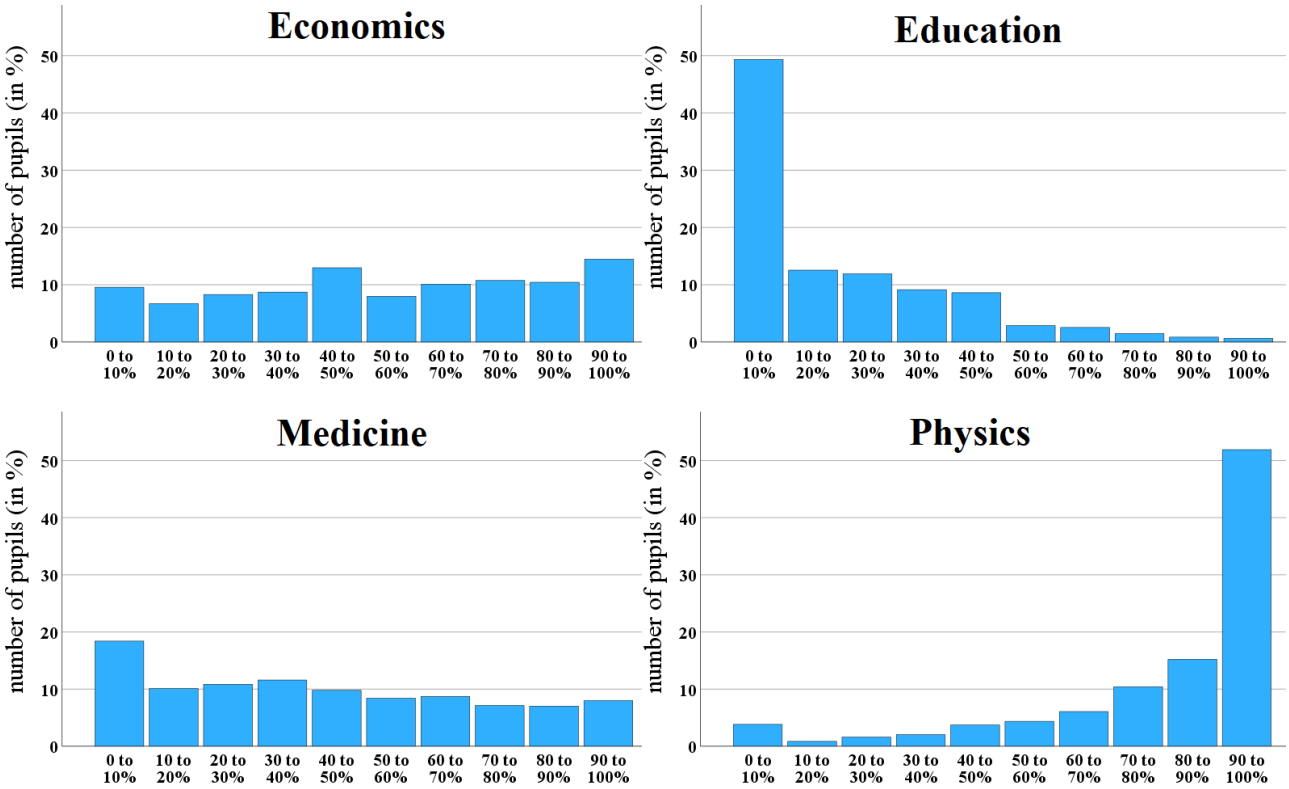


Figure 1: Distribution of students with respect to their relative frequencies of adequately scored learning prerequisites.

Table 2: Mean percentage of adequately scored learning prerequisites.

	Total	Mathematical content				Mathematical processes
		Basics	*Calculus*	*Vectors & matrices*	*Stochastics & general content*	
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*	*M (SD)*
Economics	.55 (.30)	.65 (.32)	.46 (.44)	.25 (.42)	.61 (.40)	.50 (.39)
Education	.19 (.21)	.16 (.22)	.10[b] (.29)	-[a]	.29 (.37)	.17 (.27)
Medicine	.43 (.31)	.46 (.34)	.29 (.40)	.16[c] (.33)	.48 (.43)	.46 (.39)
Physics	.81 (.25)	.86 (.24)	.79 (.36)	.75 (.39)	.68 (.40)	.83 (.31)

Notes: For each subject, only those mathematical learning prerequisites were included which had been identified as necessary by the university instructors. N = 940; MIN = 0, MAX = 1; [a] No necessary learning prerequisite; [b] Based on only one learning prerequisite; [c] Based on two learning prerequisites

Table 2 displays the mean percentage of adequately scored learning prerequisites across all participating students in more detail. Note, that these numbers are based on only those learning prerequisites which had been indicated as “necessary“ for the study of the respective subject by the university instructors. Again, the numbers reflect the quite good perceived relevance of specific mathematical prerequisites for physics, and the underestimation of prerequisites for education. Compared to basic mathematical

content (e.g. fractions and equations), which is more widely regarded as necessary, content from upper secondary level (i.e., calculus and vectors/matrices) is underestimated more often – at least for economics, medicine and physics. For the study of education, prerequisites across all topics are widely underestimated. Compared to the other content aspects, stochastics prerequisites are perceived as necessary as (or even more than) basics for the study of economics, education, and medicine. In contrast, the relevance of stochastics for physics is underestimated most often compared to the other mathematical topics. The results for the mathematical processes (i.e., processes typical of mathematical work such as modelling or argumentation), mirror the overall results.

RQ 3: Relation of adequateness of students' views to relevance of mathematics

Given the results on research questions 1 and 2, we wondered, if high school students' adequate perception of single mathematical learning prerequisites is related to their perceived overall relevance of mathematics for a particular subject of study. We therefore compared the mean percentages of adequately scored learning prerequisites between those students who had indicated no or a (rather) low relevance of mathematics in general to a particular subject (1-5 on the 10-point Likert scale) and those who had indicated a (rather) high relevance (6-10 on the 10-point Likert scale).

	No or (rather) low perceived relevance		(Rather) High perceived relevance	
	M (SD)	*N*	*M (SD)*	*N*
Economics	.33 (.27)	150	.59 (.28)	783
Education	.16 (.19)	788	.37 (.27)	144
Medicine	.28 (.25)	398	.55 (.29)	535
Physics	.48 (.41)	21	.82 (.23)	914

Table 3: Mean percentages of adequately scored learning prerequisites in two groups of students with different perceived relevance of mathematics in general for the different subjects.

Table 3 shows, that overall, high school students who perceived mathematics as not or only (rather) less relevant for the subject of study in fact rated fewer specific learning prerequisites as adequately necessary for the respective subject, compared to those students who perceived a (rather) high relevance of mathematics for the subject. For economics, a t-test showed that this difference was significant ($t(931) = -10.413$, $p < .001$) with medium effect size, $r = .32$. For the subjects of education and medicine, we performed non-parametric Mann-Whitney U tests given inhomogeneous variances and lack of normal distributions. The tests showed again significant differences with medium effect sizes for both, education ($U = 30.115$, $p < .001$, $r = .30$) and medicine ($U = 51.817$, $p < .001$, $r = .44$). For physics, the number of students perceiving no or (rather) low relevance of mathematics was too low to meaningfully perform a parametric or non-parametric test of significance. However, comparing the confidence intervals showed no overlap, indicating a significant

difference between both groups of students as well (no or (rather) low relevance: 95 % CI = [.29; .66]; (rather) high relevance: 95 % CI = [.81; .84]). Overall, these results show, that a high perceived relevance of mathematics in general is related to how adequately students perceive specific learning prerequisites for the respective subject.

DISCUSSION

Previous studies have shown that, in retrospective, many university students report that they had been surprised about how much and which mathematics is needed in their studies, sometimes resulting in unaccomplished expectations, frustration and even drop-out (e.g., Heublein, 2014; Venezia et al., 2003). Little is known, however, about in how far high school students' – prospective – perception of the relevance of mathematics for future university studies is in line with the mathematical prerequisites expected by university instructors from incoming freshmen. The present study therefore aimed at investigating high school students' views on the relevance of mathematics for future university studies, contrasting the studies of economics, education, medicine and physics.

Based on an online-questionnaire we found high school students to perceive an overall relevance of mathematics in general, but to clearly differentiate between subjects of study. Physics, a traditional STEM subject, was regarded to highly rely on mathematics. Likewise, students perceived a rather high relevance of mathematics for economics. In contrast, the relevance of mathematics for education was perceived very low. The same ranking of subjects is found when exploring students' perceived necessity of specific learning prerequisites (e.g., fractions, derivatives). However, the results on students' perceived necessity of specific learning prerequisites also shows, that students often do not know what exactly is required for studying a particular subject; this is even true for subjects, in which mathematics is considered to be highly relevant (Table 3). This finding, particularly for economics and medicine, reflects what Kollosche (2017) called an „empty signifier" (p. 639). Students view mathematics to be relevant for economics and medicine in general, but they are unaware of what aspects of mathematics are relevant in particular, and thus, how specific aspects are subjectively relevant to themselves. This relevance paradox (Niss, 1994) can only be solved if students are not only shown that mathematics is important, but also which mathematics is important. This requirement seems to be satisfactorily fulfilled for the study subject physics (Table 2) and might be due to the fact that physics instruction in German schools is permeated by mathematics. In contrast, economic classes hardly address mathematical models so that high school students know that mathematics is somehow relevant for economics but they do not have a clear idea of what mathematical content beyond the basics plays a role in economics. Finally, education is not a school subject in most of the federal states in Germany. Hence, high school students hardly have an opportunity in school to get information about the mathematical requirements for this subject of study.

Limitations

The sample selection was not representative and depended on the decisions of the school principals and teachers of the students. The participation of the students was voluntarily which might have caused a selection bias. However, since we assume that mainly interested students participated in the study, the sample might be a positive selection of students with a more positive attitude towards mathematics. A second limitation is that the mathematical requirements for the four subjects of study were determined by Delphi studies with university instructors for these subjects. Although the Delphi method is an established method to generate consensual results among the university instructors, there might be universities where study programs for the four subjects have higher or lower mathematical requirements depending on the specific focus (especially for education studies).

Implications

Despite the limitations, implications can be derived for future research as well as for educational practice. The inadequate perceptions of the relevance of mathematics to some subjects of study raises the question of what students' perceptions are based on. There might be personal factors like, for example, that students already decided to study a specific subject after finishing school and, hence, collected information about corresponding study programs. Another relevant variable might be family background that exerts influence on the students' perceptions because other family members study or studied a specific subject and possess relevant knowledge. Another type of factors might be school-related factors. With respect to the different subjects, a more mathematical orientation of the subjects' instruction in school could have an influence on students' perception of the relevance of mathematics for a specific subject of study. Such subject-specific factors could be identified by, for example, investigating the subjects' textbooks and curricula. However, since not every subject to be studied at university is represented by a corresponding school subject, contexts of application in mathematics textbooks and in mathematics class might also be analyzed in terms of both quality and quantity in order to investigate their influence on students' perception of the relevance of mathematics for subjects of study.

With respect to educational practice the present study indicates the need to show students the relevance of mathematics in various fields of study. In this context, we developed tasks for mathematics classes that highlight specific mathematical requirements in various fields of study. Initial tests with high school students have shown that students are more likely to perceive mathematical requirements in different fields of study if such tasks are supplemented by a reflection question. These reflection questions connect the content of the task with the context of university study and make the relevance of mathematics in the corresponding subject of study explicit.

REFERENCES

Brown, M., Brown, P., & Bibby, T. (2008). "I would rather die": Reasons given by 16-year-olds for not continuing their study of mathematics. *Research in Mathematics Education*,

10(1), 3–18.

Darlington, E., & Bowyer, J. (2016). Engineering undergraduates' views of A-level Mathematics and Further Mathematics as preparation for their degree. *Teaching Mathematics and its Applications*, 3, 200-216. https://doi.org/10.1093/teamat/hrw020

Deeken, C., Neumann, I., & Heinze, A. (2020). Mathematical Prerequisites for STEM Programs: What do University Instructors Expect from New STEM Undergraduates? *International Journal of Research in Undergraduate Mathematics Education*, *6*(1), 23–41.

Douglas, D., & Salzman, H. (2020). Math Counts: Major and Gender Differences in College Mathematics Coursework. *The Journal of Higher Education*, *91*(1), 84–112.

Heublein, U. (2014). Student Drop-out from German Higher Education Institutions. *European Journal of Education*, 49(4), 497–513. https://doi.org/10.1111/ejed.12097

Kollosche, D. (2017). The ideology of relevance in school mathematics. In A. Chronaki (Ed.), *Mathematics education and life at times of crisis: Proceedings of the Ninth International Mathematics Education and Society Conference* (Vol. 2, pp. 633–644). Volos, Greece: University of Thessaly Press.

Mulhern, G., & Wylie, J. (2016). Mathematical Prerequisites for Learning Statistics in Psychology: Assessing Core Skills of Numeracy and Mathematical Reasoning among Undergraduates. *Psychology Learning & Teaching*, *5*(2), 119–132.

Neumann, I., Rohenroth, D., & Heinze, A. (2021). *Studieren ohne Mathe? Welche mathematischen Lernvoraussetzungen erwarten Hochschullehrende für Studienfächer außerhalb des MINT-Bereichs? [Studying without math? What mathematical learning prerequisites do university teachers expect for non-STEM fields of study?]*. Kiel: IPN.

Niss, M. (1994). Mathematics in Society. In R. Biehler et al. (Eds.), *Didactics of Mathematics as a Scientific Discipline* (pp. 367–378). Dordrecht: Kluwer Academic Publishers

Onion, A. J. (2004). What use is maths to me? A report on the outcomes from student focus groups. *Teaching Mathematics and its Applications*, 23(4), 189–194.

Rohenroth, D., Neumann, I., Heinze, A. (2023, under review). Mathematical prerequisites for Non-STEM Programs.

Venezia, A., Kirst, M. W., & Antonio, A. L. (2003). *Betraying the College Dream: How Disconnected K-12 and Postsecondary Education Systems Undermine Student Aspirations*. Stanford, CA: Stanford Institute for Higher Education Research.

HOW ATTEMPTING TO BUILD ON STUDENT THINKING INFLUENCES THE SCAFFOLDING TEACHERS PROVIDE DURING ENACTMENT OF HIGH COGNITIVE DEMAND TASKS

Joshua M. Ruk[1] and Laura R. Van Zoest[2]

[1]University of Wisconsin-Whitewater, [2]Western Michigan University

We used videotaped enactments of high cognitive demand tasks to investigate whether teachers who were engaged in the teaching practice of building*—and thus were focused on having the class collaboratively make sense of their peers' high-leverage mathematical contributions—provided scaffolding that supported the maintenance of high cognitive demand tasks. Attempting to build on high-leverage student thinking seemed to mitigate the teachers' tendencies to provide inappropriate amounts of scaffolding because they: (1) believed the building practice required them to refrain from showing the students how to solve the task; (2) wanted to elicit student reasoning about their peer's contribution for the building practice to utilize; and (3) saw the benefits of their students being able to engage in the mathematical thinking themselves.*

INTRODUCTION

Research has shown the advantages of teaching practices that use student thinking (e.g., Carpenter & Fennema, 1992), the importance of using high cognitive demand tasks, and the need to maintain high levels of cognitive demand during task enactments to maximize their benefits (e.g., Stein & Lane, 1996). Research has also identified several factors that support the maintenance of high cognitive demand, including the appropriate use of scaffolding (Stein et al., 1996). Recent attention has been directed toward articulating specific teaching practices that use student thinking. For example, the MOST Research Team has focused on articulating the *teaching practice of building* (e.g., Leatham et al., 2022), a teaching practice designed to take full advantage of MOSTs (Mathematical Opportunities in Student Thinking)—high-leverage student mathematical contributions that provide an in-the-moment opportunity to engage the class in joint sense making about that contribution to better understand the important mathematics within it. Better understanding the interaction between teaching practices that use student thinking and the maintenance of cognitive demand will support leveraging the known abilities of both to support student learning. We contribute to this understanding by investigating how teachers' attempts to engage in the teaching practice of building affected the scaffolding they provided during their enactments of high cognitive demand tasks.

LITERATURE REVIEW

The MOST Research Team has defined the *teaching practice of building* (henceforth referred to as building) as making a student contribution "the object of consideration

2023. 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. 115-122). PME 46.

by the class in order to engage the class in making sense of that thinking to better understand an important mathematical idea" (Van Zoest et al., 2017, p. 36). They describe building as being comprised of four elements: (1) establish the student mathematics of the [contribution] as the object to be discussed; (2) grapple toss that object in a way that positions the class to make sense of it; (3) conduct a whole-class discussion that supports the students in making sense of the student mathematics of the [contribution]; and (4) make explicit the important mathematical idea from the discussion (Leatham et al., 2021, p. 1393)

Their work to articulate the key aspects of each element has relied on an iterative process involving teacher-researchers (TRs) who enacted the team's evolving conceptions of building in their classrooms (Leatham et al., 2022). The TRs did this using a set of high cognitive demand tasks designed to elicit predictable MOSTs that the TRs could prepare to build on (see Figure 1).

(a) Percent Discount	**(b) Variables**
The price of a necklace was first increased 50% and later decreased 50%. Is the final price the same as the original price? Why or why not?	Which is larger, x or $x + x$? Explain your reasoning.

Figure 1: Two tasks used by the teacher-researchers.

Cognitively demanding tasks are challenging problems, or sets of problems, that require students to use their existing knowledge, sometimes in new and unique ways, along solution pathways that are not immediately clear (Stein et al., 1996). The use of such tasks has been shown to lead to student learning gains (Stein & Lane, 1996). Unfortunately, the high cognitive demand of these tasks is not maintained in many lessons that begin with cognitively demanding tasks (Henningsen & Stein, 1997). As a result, much work has been done to understand the complexity of maintaining high levels of cognitive demand during task enactments. For example, drawing from 520 task enactments, Stein et al. (1996) identified factors that maintain and lower cognitive demand. These factors have been utilized by numerous studies (e.g., Estrella et al., 2019) to better understand the maintenance of cognitive demand.

Providing appropriate scaffolding is one of Stein et al.'s (1996) factors that help maintain cognitive demand. They determined that teachers or more capable peers offering appropriate scaffolding occurred in 58% of tasks where cognitive demand was maintained during set up and implementation. Henningsen and Stein (1997) looked specifically at tasks that began as the highest category of cognitive demand—doing math—and determined that appropriate scaffolding was offered in 73% of such tasks for which the cognitive demand was maintained. *Challenges become nonproblems* is one of the factors that Stein et al. (1996) found to cause cognitive demand to be lowered. The reasons they give for this happening included the teacher "specifying explicit procedures or steps to perform" or "either performing [the challenging aspects of the task] or telling them how to do them" (Stein et al., 1996, p. 479)—that is,

providing inappropriate scaffolding. Inappropriate scaffolding was found in 64% of enactments where cognitive demand was lowered, regardless of the starting level of the tasks (Stein et al., 1996), and in 39% of enactments where cognitive demand was lowered for doing math tasks (Henningsen and Stein, 1997). Thus, it is clear that scaffolding is an important factor in both the maintenance and decline of cognitive demand during task enactments.

Smit and colleagues (2013) conceptualized the idea of scaffolding for whole-class settings. In their work, as is typical, scaffolding is generally described as a positive contributor to student learning. The research on the maintenance and decline of cognitive demand, however, suggests that the nature of the scaffolding matters. For example, a common scaffolding strategy is to give students a worked example of a similar problem. In the context of working with a high cognitive demand task, however, this is likely to undermine the goals of the task because it provides the students with a specific strategy for solving the task, and thus lowers the cognitive demand. Teachers who are engaged in the teaching practice of building are focused on having the class engage in making sense of MOSTs—high-leverage contributions made by their peers—and thus may be focused on providing scaffolding that is more compatible with the maintenance of high cognitive demand tasks. By considering scaffolding through the lens of cognitive demand, we can draw from previous work to better understand how the scaffolding that teachers provide during task enactments may have been affected by their attempt to engage in the building practice.

THEORETICAL FRAMEWORK

Our investigation of scaffolding takes a participationist approach (Vygotsky, 1987). That is, we see student learning as taking place through students' interactions with more knowledgeable others, such as the teacher and their peers. As discussed in Sfard and Cobb (2022), this approach acknowledges several important aspects of instruction: the tasks used; the engagement of students with the tasks, both individually and through class discussion; the teacher as the knowledgeable other who facilitates the learning; and the teacher's mathematical knowledge for teaching. Thus our decision to use high cognitive demand tasks with accompanying task notes that provided guidance in these areas. For example, the notes provided common student responses to the tasks, identified which student contributions were likely to be MOSTs and provided suggested questions to ask at different points in the enactment to facilitate students' understanding of the embedded mathematics. We use a broad definition of scaffolding "as an interactional process between a person with educational intentions and a learner, aiming to support this learner's learning process by giving appropriate and temporary help" (van Oers, 2014, p. 535). Thus, we considered a teacher move to be scaffolding if there was evidence through what the teacher did or said that they were "aiming to support" the students' learning of mathematics with their actions. We used the cognitive demand framework (Stein et al., 1996) to determine whether the scaffolding provided was appropriate.

METHODOLOGY

Six middle school teacher-researchers (TRs) in the larger MOST project focused on conceptualizing the teaching practice of building (for more details, see Leatham et al., 2022) provided 24 videotaped classroom enactments of the tasks in Figure 1 (each teacher enacted each task twice), 6 online teacher surveys, and 5 recorded online teacher interviews (one teacher was not available). We first analyzed the videotaped task enactments using the **R**eorganized Factors that **U**ndermine or **K**eep Cognitive Demand (RUK; Ruk, 2020), a succinct tool designed to consistently measure the factors that maintain and lower cognitive demand (as identified by Stein et al., 1996). The "Amount of Scaffolding" category of the RUK looks at how much scaffolding was offered on a continuum from 1 to 4, with 4 representing task enactments with scaffolding that supported students without taking away necessary struggle. We next developed the survey that we administered to the TRs based on the findings from the RUK analysis. For example, to better understand scaffolding we asked this question: "In general, after you present an example or other information to your classes, how is what you presented related to the problem(s) that you assign your students afterwards?" Then we developed the interview questions for the TRs based on the results of the RUK and their individual survey responses. For example, if the results showed that a TR offered different scaffolding than usual, they were asked: "To what extent did [attempting to enact the building practice] change the amount of scaffolding that you normally offer (as opposed to other problems that you give your students)." This process of analysis followed by additional data collection allowed us to engage in the three levels of quantitative data analysis described by Simon (2019).

RESULTS

Attempting to engage in the teaching practice of building seemed to have a positive effect on the scaffolding teachers provided during their enactments of high cognitive demand tasks. The RUK showed that 23 out of the 24 task enactments (96%) had appropriate scaffolding throughout the enactment, and the remaining enactment had appropriate scaffolding at the beginning of the enactment. Contrast this with Henningsen and Stein's (1997) study of tasks of high cognitive demand, where they found appropriate scaffolding in only 73% of the enactments. To better understand the relationship of the scaffolding teachers provided to their enactment of the building practice, we turn to the interviews, where two teachers illustrated how they did a little less scaffolding than they normally do, one described how they did a lot less scaffolding, and two explained how they did the same amount of scaffolding. One of the two teachers who said that they offered a little less scaffolding gave an example of the scaffolding they would have offered for the Variables task (see Figure 1b) if they hadn't been attempting to engage in the building practice. They said that they "would have started a little bit with a conversation about what is a variable, and what different ways a variable can be used." However, this teacher recognized that offering this scaffolding "takes away from the mathematics that [the students] experienced as they went through [the task]" and, because they believed that experiencing this mathematics

was needed for the building practice, they did not offer this additional scaffolding. If the teacher had offered this scaffolding before students began grappling with this task, cognitive demand would likely have been lowered significantly. Students would have, as part of the scaffolding, discussed and likely resolved ideas related to the underlying mathematics of this task—that all possible values within a domain must be considered to determine relative values of variable expressions—before they had the opportunity to engage with these ideas themselves through exploration of the task. Thus, it seems that this teacher offered an appropriate amount of scaffolding because they believed that it was a necessary part of the building practice.

The other teacher who said that they offered a little less scaffolding, revealed during their interview that for the Variables task, if it had not been for enacting the teaching practice of building on student thinking, they "might have given [the students]: try it with two positive numbers, try it with two negative numbers, you know, can you generalize what happened." This scaffolding would have lowered the cognitive demand of this task because students would have lost the opportunity to explore for themselves and discover that positive and negative numbers (as well as zero) lead to different outcomes for this task, which in turn leads to uncovering the underlying mathematics. This teacher also said that before understanding the practice of building on student thinking, "I would have [scaffolded] right away." But now, I'd wait and "if I didn't see that there was going to be a good discussion then I might say, okay why don't you guys try it with different types of numbers and see what happens." This teacher believed that the mathematics "actually will make more sense to them because they came up with those ideas, it wasn't me telling them, oh you were wrong see here's your counterexample. You know, that they can kind of, like, work through that muddiness themselves, and come out hopefully with a more clear picture." Through attempting to engage in the building practice, this teacher came to see the value in allowing students to productively struggle in their classroom, not just in service of the building practice, but in general as well.

The teacher who said they offered a lot less scaffolding, revealed in her interview that if it had not been for enacting the teaching practice of building on student thinking, and wanting "all of those different misconceptions to come out," they would have started the Variables task by asking their students to think "about different numbers. Like, make sure you think about all the numbers, or something like that." Although this is not as explicit as telling students to consider positive and negative numbers, it would have likely had a similar effect of lowering the cognitive demand of this task because students would have been given clues about the underlying mathematics before they started grappling with the task. Thus, it also seems that the desire to draw out student thinking when enacting the building practice can improve the scaffolding offered during the use of high cognitive demand tasks.

One of the two teachers who said they would have offered the same amount of scaffolding offered an appropriate amount of scaffolding during their enactments. This teacher had an extensive knowledge of cognitive demand research, and said that this

research, “made a huge impact on me, and my practice” and that “I think it's extremely important that students are engaged and doing significant math and working at a high level.” This teacher was immersed in applying cognitive demand research in their classroom and thus was already focused on continuously attempting to provide an appropriate amount of scaffolding.

The other teacher offered inappropriate scaffolding in one of their enactments by working through a specific example that broke the problem down into smaller steps, and then asking students for only small bits of information, such as simplification of expressions or what specific numbers represented—a type of questioning pattern that Wood (1998) identified as funneling. Also, rather than letting students work through incorrect ideas by asking follow-up questions, the teacher said “no,” and waited for correct thinking to emerge. These actions run contrary to the practice of building because, although the interaction began with the student’s thinking, the attention quickly shifted to the teacher’s way of thinking. These teacher actions removed the challenge from the task by breaking it down into smaller parts, controlling the conversation, and only moving forward when students shared the correct thinking that the teacher was looking for. Fortunately, the teacher did this towards the end of the enactment, so the students had time to grapple with the task before the cognitive demand was lowered. Had the teacher provided this scaffolding earlier in the enactment, cognitive demand likely would have been lowered even more. To gain an understanding of why cognitive demand was lowered in this way, we look again at our interview data.

During their interview, this teacher realized that offering the scaffolding that they did ran contrary to the building practice—that they had gone too far and had given too much information away to their students. The teacher also noted that if they had not been enacting the practice of building on students’ thinking, they likely would have worked through an example like this much earlier in the class discussion but held off because they believed that enacting the building practice required more time for students to work through this task on their own. Furthermore, if it hadn't been for enacting the building practice, she likely would have lowered the cognitive demand even further than she did by offering this scaffolding earlier and giving her students even less time to productively struggle with this task. So, even though she did not adhere to the guidelines of the practice, this teacher still maintained a higher level of cognitive demand than if she had not been attempting to enact the building practice.

Considering the cases described above, the results of applying the RUK, and the survey and interview data, we can conclude that the teachers in this study were able to recognize appropriate scaffolding for enacting a high cognitive demand task. However, even though they can recognize this, they may still offer scaffolding that takes away the need for students to make sense of the mathematics in the task, and thus lowers the cognitive demand of the tasks that they are enacting. Attempting to build on student thinking seemed to mitigate the teachers’ tendencies to provide inappropriate amounts of scaffolding for three reasons: (1) they believed the building practice required them

to hold back from showing the students how to solve the task: (2) they wanted to elicit student reasoning about their peer's contribution for the building practice to utilize; and (3) they saw the benefits of their students being able to engage in the mathematical thinking themselves.

DISCUSSION

Our findings showed that the teachers in this study were able to recognize appropriate scaffolding. Indeed, almost unanimously they provided appropriate scaffolding for every task they enacted as part of this study. However, even though they could all recognize, and successfully provide appropriate scaffolding, most of them noted that they would have offered more scaffolding had they not been enacting the building practice. Furthermore, if they were all to provide the scaffolding that they described in their interviews, they almost assuredly would have lowered cognitive demand for their task enactments. However, it seems that if teachers have a specific reason (e.g., attempting to enact the building practice, or a belief in the importance of maintaining cognitive demand), they can, and do, provide scaffolding that supports student learning.

Since teachers can recognize appropriate scaffolding and provide it if they try, it seems that what they need is a reason to do so. In our study, attempting to engage in the building practice provided that reason. Other specific teaching practices may also provide reasons for offering appropriate scaffolding and thus support the use of scaffolding to maintain high cognitive domain. As such, future research to better understand the influence of specific teaching practices on the scaffolding teachers provide during enactment of high cognitive demand tasks could compare teachers who simply enact a task with teachers who try to engage in a teaching practice that prompts them to consider the scaffolding they should offer when enacting that same task. Future research could also investigate whether teachers who have improved the scaffolding they provide when engaged in a particular teaching practice, such as building, extend that appropriate use of scaffolding to their teaching more broadly.

ACKNOWLEDGEMENTS

This research report is based on work supported by the Western Michigan University Dissertation Completion Fellowship and by the U.S. National Science Foundation (NSF) under Grant Nos. DRL-1720410, DRL-1720566, and DRL-1720613. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

REFERENCES

Carpenter, T., & Fennema, E. (1992). Cognitively guided instruction: Building on the knowledge of students and teachers. *International Journal of Educational Research, 17*, 457-470. http://dx.doi.org/10.1016/s0883-0355(05)80005-9

Estrella, S., Zakaryan, D., Olfos, R., & Espinoza, G. (2019). How teachers learn to maintain the cognitive demand of tasks through lesson study. *Journal of Mathematics Teacher Education.* Advanced online publication. doi: 10.1007/s10857-018-09423-y

Henningsen, M., & Stein, M. K. (1997). Mathematical tasks and student cognition: Classroom-based factors that support and inhibit high-level mathematical thinking and reasoning. *Journal for Research in Mathematics Education, 28*, 524-549. doi:10.2307/749690

Leatham, K. R., Van Zoest, L. R., Peterson, B. E., & Stockero, S. L. (2022). A Decomposition of the Teaching Practice of Building. In *Proceedings of the NCTM 2022 Research Conference*.

Ruk, J. M. (2020) A conceptualized framework for assessing factors that maintain and lower cognitive demand during task enactment. In A. I. Sacristán, J. C. Cortés-Zavala, & P. M. Ruiz-Arias (Eds.), *Proceedings of the 42nd Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 2351-2352). Mazatlán, Mexico.

Sfard, A., & Cobb, P. (2022). Research in Mathematics Education: What Can It Teach Us about Human Learning? In R. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (Cambridge Handbooks in Psychology, pp. 467-485). Cambridge University Press. doi:10.1017/9781108888295.029

Simon, M. A. (2019). Analyzing Qualitative Data in Mathematics Education. In K. Leatham (Ed.), *Designing, Conducting, and Publishing Quality Research in Mathematics Education* (pp. 111-122). Switzerland AG: Springer.

Smit, J., van Eerde, H. A. A., & Bakker, A. (2013). A conceptualisation of whole-class scaffolding. *British Educational Research Journal, 39*(5), 817–834.

Smith, M. S., & Stein, M. K. (2018). *5 practices for orchestrating productive mathematics discussions* (2nd ed.). Reston, VA: National Council of Teachers of Mathematics.

Stein, M. K., Grover, B. W., & Henningsen, M. (1996). Building student capacity for mathematical thinking and reasoning: An analysis of mathematical tasks used in reform classrooms. *American Educational Research Journal, 33*, 455-488. doi:10.2307/1163292

Stein, M. K. & Lane, S. (1996). Instructional Tasks and the development of student capacity to think and reason: An analysis of the relationship between teaching and learning in a reform mathematics project. *Educational Research and Evaluation, 2*(1), 50-80. Doi: 10.1080/1380361960020103

van Oers, B. (2014). Scaffolding in Mathematics Education. In S. Lerman (Ed.), *Encyclopedia of Mathematics Education*. Springer, Dordrecht.

Van Zoest, L. R., Stockero, S. L., Leatham, K. R., Peterson, B. E., Atanga, N. A., & Ochieng, M. A. (2017). Attributes of instances of student mathematical thinking that are worth building on in whole-class discussion. *Mathematical Thinking and Learning*, *19*(1), 33-54. https://doi.org/10.1080/10986065.2017.1259786

Vygotsky, L. S. (1987). Thinking and speech. In R. W. Rieber & A. C. Carton (Eds.), *The collected works of L. S. Vygotsky* (pp. 39-285). New York: Plenum Press.

Wood, T. (1998). Alternative patterns of communication in mathematics classroom: Funneling or focusing? In H. Steinbring, M. G. Bartolini Bussi, & A. Sierpinska (Eds.), *Language and communication in the mathematics classroom* (pp. 167-178). National Council of Teachers of Mathematics.

TEACHER CHANGE AND INCLUSIVE INTERVENTIONS FOR LEARNERS WITH MATHEMATICS DIFFICULTIES

Robyn Ruttenberg-Rozen and Marc Husband

Ontario Tech University and St. Francis Xavier University

Inclusion for learners experiencing difficulties learning mathematics (DLM) relies partly on the actions of the teacher in the anticipating, planning and teaching of interventions. Yet, teacher enactment of mathematics interventions can be variable and bereft of mathematics. We share findings on how 16 teachers experienced foundational changes to their conceptions of interventions. Teachers participated in a video club where we educated awareness of inclusive mathematical interventions for learners experiencing DLM. Teachers reconceptualized their role as being embedded within interventions, responsive, and agentic. Findings contribute to the understanding of teacher change that supports inclusive practice for learners experiencing DLM.

INTRODUCTION

The promise of an inclusive mathematics classroom remains elusive, despite continued calls and research (Qvortup & Qvortup, 2018). Streaming, ability grouping, or removing students from the classroom according to perceived math ability and mathematics difficulty are still prevalent in many mathematics classrooms (e.g., Hunter et al, 2020). Judgements of math ability and difficulties are tied to biases and beliefs about who can do mathematics (Ruttenberg-Rozen & Jacobs, 2022), leading to tiered systems where some students are provided with a more rigorous mathematics education than others. Inclusion is a complex construct that can be understood in relation to exclusion, and who is excluded. While there are many definitions for inclusion, we use Qvortup and Qvortup's (2018) comprehensive definition of inclusion as a three-level process of (1) physical inclusion, (2) social inclusion, and (3) identity (belonging to a community).

All three processes rely on the actions of the teacher in the anticipating, planning, and teaching of interventions for learners experiencing difficulties learning mathematics. However, the practice of these processes competes with deficit ability narratives about who can learn mathematics. These narratives can be unconscious and differ from conscious beliefs. For example, a teacher might hold the conscious belief that everyone can learn and be successful at learning mathematics, but simultaneously have deficit narratives that students who experience DLM cannot be successful at mathematics. This means that despite conscious beliefs supporting inclusion, teachers may not actively support the physical, social, and identity inclusion of learners who experience DLM because of deficit narratives (Ruttenberg-Rozen & Jacobs, 2022).

The goals of the current study are to use Mason's (2002) Educating Awareness to provoke teacher change towards more mathematical and inclusive interventions for learners experiencing DLM. In this study we ask the questions:

2023. 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. 123-130). PME 46.

- How do teachers conceptualize mathematics interventions and their role in interventions for learners experiencing difficulties learning mathematics?
- What inclusive shifts do teachers make in their conceptions of interventions through developing Awareness of narratives of practice?

We note that the term "mathematics difficulties" is used in many ways in the research literature. In this present study we asked the teachers to identify learners who they perceived as having difficulty learning mathematics. The teachers chose students whom they felt were "unresponsive" to their teaching.

INTERVENTIONS

The term "intervention" is often used as a catch-all phrase in mathematics classrooms to encompass any action that involves intervening between learners experiencing difficulty and mathematics content (Fuchs et al., 2008). Despite a large number of documents, professional development, and guides provided to teachers about practice for interventions, research has found that teachers rarely enact recommended strategies in the classroom (Crawford et al., 2020). When teachers do enact the strategies, there is still variability even with pre-designed intervention programs (Hunt et al., 2016). How teachers conceive of an intervention depends on prior experience and perspective with interventions. For instance, an intervention in mathematics can be conceived of as something planned for after an assessment, when the teacher sends students to a different teacher for support, or in-the-moment responses that teachers enact. Notably, the content of mathematics interventions for learners experiencing mathematics difficulties is often bereft of mathematical reasoning, despite the interventions being aimed at mathematics learning (Gervasoni & Lindenskov, 2011). For example, in the practice of explicit feedback in mathematics interventions (Powell & Fuchs, 2015, p.185), teachers use "affirmative" feedback, noting the procedures the student did well, or "corrective" feedback, "redirecting all student errors." Although both types of feedback serve the purpose of supporting learners in meeting the requirements of certain curricular goals, they do not necessarily support mathematical reasoning.

Subsequently, it is important for teachers to develop a practice of mathematical interventions that is inclusive and mathematical. In this study we adapt Qvortup and Qvortup's (2018) definition of inclusion to mathematics interventions. For mathematics interventions to be inclusive they should:

- Support the physical needs of learners by providing adequate space to do and play with mathematics, with all learners,
- Support the social needs of learners by acknowledging the intervention of mathematics within the classroom "as a complex of interaction systems" (p.812), leveraging the systems in an anti-deficit way for mathematics learning, and
- Support the mathematical identity of learners by providing support to both achieve mathematically and develop the mathematical reasoning skills that supports being part of a mathematical community.

THEORETICAL FRAMEWORK—FOUNDATIONAL TEACHER CHANGE

In the literature on teacher change, researchers have proposed different models for looking at what and how teachers change (e.g., Guskey, 2002). Although these models attribute different factors (e.g., dispositions, practices, beliefs) to view changes and discuss different pathways for change (e.g., linear vs cyclical), there are common underlying experiences that support teacher change, including an experimentation phase where teachers try things out in their own classrooms (Putnam and Borko, 2000) and the value of discussing and collaborating with colleagues (Goldsmith et al., 2014). These experiences are offered to teachers through a range of professional development opportunities, including structured courses with timelines and self (teacher) directed inquiries that occur without imposed timelines or expectations. For example, Chapman and Heater's (2010) study documents one teacher's self-development journey throughout her career, characterizing awareness and teacher change through three types or processes: instrumental, conceptual, and foundational. Of significance to this study is foundational change where Chapman and Heater (2010) describe the change from both the teacher's orientation of self (identity) and practice (actions). Teachers embracing foundational change do so by thinking in new ways and are empowered to transform their teaching.

THEORETICAL FRAMEWORK—EDUCATING AWARENESS

An awareness is a concept that encapsulates the conscious and unconscious sight (Mason, 2002) we gain from our noticings about our experiences. The path to an awareness begins with a disturbance of some kind that triggers a noticing,

> for the disturbance which triggers a noticing triggers a collection of associated sensitivities, and hence also triggers a perspective, a way of seeing and of thinking about what is noticed (Mason, 2002, p.167).

The disturbance challenges the familiar and causes our senses to be heightened. We then direct our senses to attend to something within the disturbance, creating a noticing (Mason, 1998). What we attend to and the resultant noticing depends on our individual experiences, expertise, and backgrounds. As such, the same disturbance can create different noticings for each person. For an awareness to occur, there has to be "intention and commitment" (Mason, 2002, p.36) at each attention and noticing phase. A person's degree of awareness and amount of intention and commitment are influenced by expertise. Experts become accustomed to their environments and may not recognize the disturbance in the first place. At the same time, experts can leverage their multiple connected experiences for interpretation (Mason,1998) at the noticing and awareness phases. Noticings are responsive to events once those events have already occurred. However, the goal of an awareness is anticipatory so that we can move closer and closer to notice the moment before an event, the moment when one still has a choice to make (Mason, 2002).

Educating awareness (Mason, 2002; Mason & Davis, 2013) is about educating for explicitly constructing a repository of awareness experiences. The purpose of the

repository is to connect the experiences for meaning making at the moment of noticing. Essentially, educating awareness is developing the expertise for the many potential opportunities that can result from one noticing.

METHODS

This research project used video club as a method for professional development within a 36-hour course about assessment and elementary mathematics education. Much literature has been written about using video clubs to frame professional development in mathematics education (van Es & Sherin, 2008). This includes how video clubs can support teacher practice and reflection (Charalambos et al., 2018). There were 21 elementary (grades 1-8) teacher participants in this study with a range of experience from 1 to 15 years. Eight teachers signed up to share video clips (30 seconds to 2 minutes) from their classrooms that evidenced an intervention. The video clubs took place during one semester, and the clips came from various elementary school grades across a large geographical area that included both rural and urban school settings. After the course was completed, teachers were asked to complete a short reflection based on the following prompt: *What new awareness(s) do you have about interventions? If you can remember, what were you doing when you first had that new awareness. I used to think that interventions were …… Now I know interventions are…*

We used thematic analysis (Braun & Clarke, 2006) to analyze our data. We first coded all the reflections using Concept Codes, "a word or short phrase that symbolically represents a suggested meaning broader than a single item or action" (Saldana, 2016, p.119). We coded three types of responses: original conceptions of interventions, new shared awarenesses, and evidence of change. We, then collapsed all the codes into our three identified themes: embedded shifts, responsive shifts, and agentic shifts.

RESULTS

We found evidence of foundational changes regarding shifts of awareness of interventions as inclusive in 16 of our 21 participants. Of the other five participants, three had conceptual change only, and we did not find evidence of change in the last two participants. The 16 participants had varying conceptions of interventions as their starting points (e.g., interventions are what specialists do, and interventions as rote practice). However, regardless of their starting point, each of the 16 participants developed new awarenesses in at least one of our three themes. Table 1 shows the distribution of themes among the16 participants. In what follows we share our results regarding the three themes. For each theme, we highlight one reflection that demonstrates shifts in awareness.

Table 1: Distribution of Themes

Themes	Participants
Embedded	Nadine, Shanti, Grayson, Greta, Seshi, April, Talia, Mariam, Sheryl
Responsive	Nadine, Grayson, Malika, K.T., Greta, Kathy, Tracy, Sheryl, Nkechi, Pat
Agentic	Nadine, Shanti, Tariq, Greta, Kathy, Sheryl, Nkechi

Embedded Shifts

We found 9 instances of *embedded shifts* in our analysis of the participants' reflections. We defined a shift as embedded if participants demonstrated a new awareness of interventions as an explicit part of their mathematics curriculum and lesson planning, and/or as embedded throughout the lesson. Some of the participants, like Greta, originally thought that "Intervention(s) was something I really thought…the Resource Teacher…did." Other participants, like Mariam, originally viewed interventions as something that only happened at the end of a lesson after a student was assessed. The participants who experienced embedded shifts discussed how they could plan mathematics tasks and teacher moves to be used as interventions. Nadine, for example, highlighted how she could use mathematical tasks and plan questions to support the needs of her learners experiencing difficulties:

> Questions that are well designed…get all students doing math. (they can be as simple or as complex as you make them). I think that simple interventions like providing students with rich tasks and thinking questions will give you the most out of your precious teacher time.

Before participating in video club, Nadine described how she prepared separate lessons for her 'struggling students.' These lessons would be implemented after the whole group lesson. Nadine has now shifted her awareness of interventions to be embedded (included) within her regular teaching practice. Nadine's reflection demonstrates a foundational change in how she now is considering how to implement (action) her new awareness within the time constraints of her teaching practice.

Responsive Shifts

We found 10 instances of *responsive shifts* in our analysis of participants' reflections. We defined a shift as responsive if participants demonstrated a new awareness of interventions as being interactional (e.g., not something one does to another), and in-the-moment when a teacher notices a difficulty. Some of the participants, like Malika, originally thought that interventions were only "to gather evidence" and assess the learners. Other participants, like Cheryl, saw interventions as "scripted" and "proscribed" for teachers to carry out. The participants who experienced responsive shifts discussed in-the-moment mathematical conversations between learner and teacher or between learners as being interventions for learners who experience

mathematics difficulties. Grayson, for example, highlighted how the conversation in random groups during rich mathematical tasks could be an intervention,

> Students who are given rich math tasks are provided an opportunity to contribute meaningfully in randomized groups. These groups are not based on what we perceive as abilities of the student, but on what developing conversation they can have about the material. These groups lead to deeper understanding of mathematical discourse.

Before participating in video club, Grayson thought that interventions were only what were proscribed on an Individual Education Plan (IEP) and were solely planned by the resource teacher. Grayson has now shifted his awareness to the possibility of randomized groupings, and students intervening with each other. Notably, Grayson now includes learning about "mathematics discourse" as an intended outcome of an intervention. This is a foundational change for Grayson as he begins to include mathematical thinking processes in his concept of an intervention.

Agentic Shifts

We found seven instances of *agentic shifts* in our analysis of participants' reflections. We defined a shift as agentic if participants demonstrated a new awareness of their power to provide inclusive interventions for their students. Some of the participants discussed interventions as outside their purview and their sense of professional failure when deferring to others for interventions. Tariq was the most explicit of all our participants in this regard, when he shared,

> …interventions meant the last resort. After you try everything you can on your own … you contact the intervention specialist. This expert arrives at your classroom to make an hour-long observation and will give you suggestions to fix whatever is it you called them about... Oftentimes reaching out for this help makes a teacher feel like they are failing their students in some way.

Tariq shares his new awareness,

> I have learned that interventions are constant, and a result of teacher noticings. The second I notice something a student has said, or has done, and I make a note for future instruction, or ask a question, or back away and give students time to think, I am making interventions to direct, or redirect thinking and deepen student understanding.

Tariq begins to express agency in his narrative of new awarenesses about interventions. Whereas before Tariq saw interventions as something an outside expert enacts, he now realizes that he can enact interventions with his teacher moves. Added to this Tariq is empowered by the notion that he is already doing these interventions in his classroom. Other participants also use agentic words to describe their new awarenesses, like Kathy who discusses "the meaningful impact" she will make through responding to students, and Shanti who speaks about interventions as "something I can do."

DISCUSSION AND CONCLUSION

In our study, we provided 21 participants with noticing opportunities to develop new awarenesses of inclusive mathematical interventions. 16 of the 21 participants experienced foundational change. This means that 16 participants developed a new orientation in their understanding of interventions and discussed the subsequent changes they would make to their practice. Regardless of their starting point, participants experienced change in at least one of three ways: a shift towards viewing interventions as embedded, a shift towards viewing interventions as responsive, and/or a shift towards regarding their own agency to implement interventions in their practice. Taken together, these shifts correspond to the expanded definition (Qvortup & Qvortup, 2018) of inclusive mathematics interventions we included above.

Teachers became more aware of how interventions can be enacted within the physical space of their classrooms. Meaning students who experience difficulties learning mathematics can experience physical inclusion through the interventions. In terms of social implication, teachers became more aware of how to leverage their learning communities to support interventions for learners experiencing difficulties. In these learning communities all learners, regardless of difficulty could receive supports and suggestions. Finally, and perhaps most significantly, the foundational changes that participants experienced has the potential to impact the mathematical identities of their students. Many of the mathematics interventions currently in practice are not mathematical (Gervasoni & Lindenskov, 2011). Through the video club, the teachers became aware that mathematics interventions include mathematics and the mathematical thinking processes. Mathematics is embedded, responsive, and teachers felt agency to include them in interventions. Integrating mathematics and mathematics thinking processes into their intervention practice, can support the sense of belonging and entry into a mathematics community for learners experiencing difficulties.

ACKNOWLEDGEMENTS

We acknowledge the generous funding of this research by the Social Sciences and Humanities Research Council of Canada

REFERENCES

Ball, D. L., & Bass, H. (2003). Making mathematics reasonable in school. *A research companion to principles and standards for school mathematics*, 27-44.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101.

Crawford, A. R., Johnson, E. S., Zheng, Y. Z., & Moylan, L. A. (2020). Developing an understanding procedures observation rubric for mathematics intervention teachers. School Science and Mathematics, 120(3), 153-164.

Chapman, O., & Heater, B. (2010). Understanding change through a high school mathematics teacher's journey to inquiry-based teaching. *Journal of Mathematics Teacher Education*, *13*(6), 445-458.

Charalambous, C. Y., Philippou, S., & Olympiou, G. (2018). Reconsidering the use of video clubs for student-teachers' learning during field placement: Lessons drawn from a longitudinal multiple case study. *Teaching and Teacher Education*, *74*, 49-61.

Fuchs, L. S., Fuchs, D., Powell, S. R., Seethaler, P. M., Cirino, P. T., & Fletcher, J. M. (2008). Intensive intervention for students with mathematics disabilities: Seven principles of effective practice. *Learning Disability Quarterly*, 31(2), 79-92.

Gervasoni, A., & Lindenskov, L. (2011). Students with 'special rights' for mathematics education. In B. Atweh, M. Graven, W. Secada, & P. Valero (Eds.) Mapping equity and quality in mathematics education (pp. 307-323). Springer.

Goldsmith, L. T., Doerr, H. M., & Lewis, C. C. (2014). Mathematics teachers' learning: A conceptual framework and synthesis of research. *Journal of Mathematics Teacher Education*, *17*(1), 5-36.

Guskey, T. R. (2002). Professional development and teacher change. *Teachers and Teaching*, *8*(3), 381-391.

Hunt, J. H., Valentine, C., Bryant, D. P., Pfannenstiel, K. H., & Bryant, B. R. (2016). Supplemental mathematics intervention: How and why special educators intensify intervention for students with learning disabilities. *Remedial and Special Education*, 37(2), 78-88.

Hunter, J., Hunter, R., & Anthony, G. (2020). Shifting towards equity: challenging teacher views about student capability in mathematics. *Mathematics Education Research Journal*, 32(1), 37-55.

Mason, J. (1998). Enabling teachers to be real teachers: necessary levels of awareness and structure of attention. Journal of Mathematics Teacher Education, 1, 243-267.

Mason, J. (2002). Researching your own practice: The discipline of noticing. Routledge Falmer.

Mason, J., & Davis, B. (2013). The importance of teachers' mathematical awareness for in-the-moment pedagogy. *Canadian Journal of Science, Mathematics and Technology Education*, *13*(2), 182-197.

Powell, S. R., & Fuchs, L. S. (2015). Intensive intervention in mathematics. *Learning Disabilities Research & Practice*, 30(4), 182-192.

Putnam, R. T., & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning?. *Educational Researcher*, *29*(1), 4-15.

Qvortrup, A., & Qvortrup, L. (2018). Inclusion: Dimensions of inclusion in education. *International Journal of Inclusive Education*, 22(7), 803-817.

Ruttenberg-Rozen, R., & Jacobs, B (2022). Considerations of Equity: At the nexus between mathematics and special education. In Xin, Y.P., Tzur, R., & Thouless, H. (Eds.) *Enabling mathematics learning of struggling students.* (pp.3-18). Springer.

Saldaña, J. (2016). The coding manual for qualitative researchers (3rd ed.). Sage.

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.

WHAT IS A "GOOD" ARGUMENTATION IN MATHEMATICS CLASSROOM?

Saccoletto Marta

University of Turin

This paper investigates the argumentative processes produced by upper secondary school students during an educational activity designed with the dual goals of introducing classical concepts of probability theory through problem solving and promoting argumentative competence. We reckon that in this context the main function of argumentations is to support probabilistic thinking development. We are interested in investigating how this function influences the development of argumentative processes in the classroom. We analyse two classroom episodes in order to show how argumentative processes evolve in relation to the interventions of the teacher and the peers, and how the changes in the produced arguments reflect the function that argumentation is intended to serve.

INTRODUCTION

In the educational context, the teaching and learning of argumentation is receiving increasing attention and it is included in the official documents of several countries. In Italy, argumentation appears within the learning goals at every scholastic level, even with explicit reference to mathematics (Mariotti, 2022). A reflection on the aspects of argumentation that can be considered and assessed in the mathematical classroom is therefore necessary (Stylianides et al., 2016). This depends on the conceptualisation of argumentation. Many Mathematics Education researchers have investigated the topic of argumentation from different perspectives and there is no single definition of argumentation in the field (Hanna, 2020). As far as mathematics education is concerned, some research highlights the social side of argumentation and focuses on the aspects of the argumentation that influencing acceptance by the classroom. Stylianides (2007) notices that the statements used in an argument should be in line with certain standards of the current mathematical culture and, at the same time, it should be accepted within the conceptual reach of the classroom participants. Classmates and especially the teacher play a crucial role. Interlocutors – and therefore their beliefs, knowledge, and convictions – can influence the development of the argumentative process (Krummehuer, 1995) and the constitution of socio-mathematical norms (Yackel, 2001). Students-teacher cooperation is usually exercised through the use of language, which plays a key role in mathematics classroom, and the development of (written and/or spoken) texts (Ferrari, 2021). Argumentation processes can be developed in classrooms in relation to different activities and purposes (not only to convince). For example, argumentation can be taken into consideration in relation to conjecture (Pedemonte, 2007), or it could be used to make students' thinking visible (Cusi et al., 2017). In addition, it is often related to students' learning (Schwarz, 2009).

2023. 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. 131-138). PME 46.

We think that the different functions of the argumentative process in classrooms are related to the educational goals. This work is part of the research strand that considers argumentation primarily as a social process, based on production of texts. We focus on argumentations developed during an upper secondary school classroom sequence, designed with the double aim of introducing classical probability theory concepts through problem-solving and fostering argumentative competence. In this context, the development of argumentative processes and probabilistic thinking are interconnected. In particular, argumentations sustain the development of sensemaking in probability. The core of the paper is investigating how the argumentation function of sustaining learning influences the development of argumentative processes in the classroom.

THEORETICAL FRAMEWORK

In line with Ferrari (2019), we consider argumentation as an interactive process based on the production of (written and/or spoken) texts. Texts and, more generally, language play different functions in mathematical classroom. For example, Ferrari (2021) distinguishes functions that are related to everyday use and functions typical of mathematics education. His research makes use of tools of linguistic pragmatics, whose main interest is the study of different uses of language. As far as argumentation is concerned, the neuropsychologists Mercier and Sperber (2017) thoroughly study uses of argumentations and reasoning. In accordance with the authors, arguments are mostly intended for social consumption. People may present them to explain and justify themselves, to evaluate others' reasoning, or to convince those who think differently. Producing and sharing arguments generally make communication more advantageous, by making it more reliable. A good argumentation should display coherence relationships between the speaker's claim and the knowledge, convictions and system of beliefs held by the addressees, allowing the audience to evaluate these relationships on their own. The ways in which the coherence relationship can be displayed and evaluated depend on the context and on cultural aspects. Moreover, in Mercier and Sperber's (2017) theory, the role of the interlocutor/s is fundamental. As in the broader case of communication, the interactive nature of the dialogue and the interlocutor's responses allow to refine justifications and arguments, shaping the argumentation process. The interlocutor's reactions are particularly useful for two reasons: she/he can indicate whether she/he has understood, and she/he can actively guide the effort of the speaker. In our study we mainly focus on conversations developed in educational settings that are characterised by educational aims. We are hence interested in reactions that contain information about student's performance and understanding in mathematical classroom. In accordance with the framework of Hattie and Timperley (2007) we call such reactions feedback. The main purpose of feedback is to help learners shorten the distance between current understanding and performance and the intended educational goals. The effective feedback addresses the issues of clarifying the direction of the teaching and learning trajectory, of considering the progress that has been made towards the goal and of defining which activities are useful to make better progress. It could address multiple issues at once. Feedback works at four levels:

Task level (in relation to quality of the task implementation), Process level (concerning the process underlying a task), Self-regulation level (referring to the way students face the task and the achievement of the learning goals, including control and confidence), Self level (usually expressed as praise addressed to students).

In line with the theoretical framework and the context of the didactical sequence considered, we focus on argumentative processes mainly addressed to other classmates and responding to the function of supporting the learning processes occurring in the mathematics classroom. The question the article seeks to answer is: how do these argumentative processes evolve in relation to teacher and student feedback, and which ways to express the coherence relationships between the speaker's statement and the knowledge and belief system of the peers emerge?

METHODOLOGY

We collected data from a didactical sequence that was designed with the teacher and that was implemented during regular mathematics classroom activities of 19 students of 11th grade. Two main educational goals were involved: to introduce classical probability theory concepts through problem-solving and to foster collaborative argumentative processes. The last one is at the core of the paper. The didactical sequence design included two main phases. In the first phase students were engaged with introductory activities focused on the resolution of the classical problem of “division of the stakes”, historically associated with the emergence of classical probability concepts in the Pascal-Fermat correspondence (Borovcnik & Kapadia, 2014). In the second phase, students were challenged with different problems, which allowed them to face some of the misconceptions typical of the field (Batanero, 2005). In this work we consider only argumentative processes developed during the first phase. This phase focused on the “division of the stakes” problem, which is:

Two players A and B play heads or tails with a fair coin. Each game, corresponding to each coin toss, is won by A if the outcome of the toss is heads and by B if the outcome is tails. A and B give 12 euros each. The stake is 24 euros. The player who first wins 6 rounds wins the game, and thus the entire stake. A always bets on "heads" and B on “tails”. The game is interrupted at the score 1- 0 for A. How should the stakes be fairly divided i.e., that it gets both players to agree?

The problem was presented to the students before any theoretical probability concept related to it. Firstly, students faced the problem in small groups and shared their resolution to the whole class (step 1). Secondly, students returned in their small group to answer some teacher’s questions that help them to analyse their and other groups’ resolutions with a critical stance. The reflections were then shared during a collective discussion (step 2). Subsequently students were presented with some solutions given by mathematicians in the history of mathematics. We chose four resolutions: Pacioli’s, Cattaneo’s, Cardano’s and Fermat’s. Mathematicians’ solutions were proposed for the problem in which the game was supposed to stop at a different score: 5- 3 for A. For the detail of the resolution and the ways in which they have been presented to students

we refer to Paola (2019). Each of the four students' groups received a different resolution. Students then worked in small groups to understand the proposed resolution and they presented it at the rest of the classroom (step 3). Afterwards, in small groups' students reflected and critically analysed the proposed resolutions. Finally, during class discussion, students were asked to argue in order to sustain or to reject the proposed resolutions (step 4).

The lessons were video recorded, and the students' written productions were collected. Interactive argumentations have been the focus of a qualitative analysis.

PRESENTATION AND ANALYSIS OF TWO CLASSROOM'S EPISODES

In this section, we present and analyse data from the last two class discussions. Two episodes are described, in which students from the same group (G1) propose different arguments to support the same conclusion, and the related reactions of peers and the teacher. In addition, we briefly describe two representations introduced by another group (G2), that are crucial to analyse the second episode.

Episode 1

The first episode occurs during the class discussion in step 3. G1 study the mathematician's solution that was assigned to them and then present it to the classroom. Since they do not agree with the solution, they show another resolution they came up with. Below, we report Andrea's presentation of the G1 proposal and the subsequent reaction of one of his classmates.

Andrea: We calculated that, now I'll explain how, A has a chance of winning by 7/8 while B by 1/8. Why? Because we start from a situation of 5 to 3. And if we assume […] that B wins, tails would have to come up three times in a row. However, the probability that of getting tails three times in a row is one over two cubed. Why? Because the probability of getting heads or tails is always fifty percent, so one-half, and if this must happen three times in a row it will be one-eighth and therefore two cubed. Consequently, the probability of A winning is seven-eighths…

Enrico: I don't understand why, if it has to come three times heads, I have to multiply three times a half.

At this point in the discussion almost all students agree to divide the stakes proportionally to the probability of the two players winning. However, as declared by Enrico, it is not clear how to calculate these "probabilities of winning". Andrea's text is mathematically correct, and Andrea seems to be willing to explain and justify each step of his group solution. However, his text is based on some assumptions about probability computation that are not shared in the classroom. Enrico's question highlights this lack of understanding and agreement with the underlying assumption. Since Andrea is not able to explain why the result was 1/8, another G1-participant, Giacomo, tries to help. Giacomo is the only one in the classroom who had already encountered probability theory in previous scholastic segments. Firstly, Giacomo tries to explain the result showing all the possible outcomes of three tosses of a fair coin.

However, this does not explain why they used the multiplication, so Enrico raises the issue again.

Enrico: However, I don't understand why the product…

Giacomo: Because at the algebraic level if you consider the probability of coming out tails to be equal to one half, if you toss the second coin, this [toss] is independent of flipping the first coin. Being two independent events, you have to multiply one half by one half.

Enrico: But why? I don't understand.

Giacomo's justification is based on the knowledge of some classical probability computational algorithms, which are not known to the other students in the classroom. While Andrea's argument does not explain why the result of the calculus is 1/8, in Giacomo's argument the justification is presented, but it is based on theory, which seems to be evoked in an almost authoritative way. According to Mercier and Sperber's (2017) definition, both Andrea's and Giacomo's interventions are not good argumentations, since their peers are not able to evaluate the coherence relationship between the speaker's claim and their knowledge, convictions and system of beliefs. At this point, the teacher decides to stop the discussion. He states that Giacomo's assertion about the product of stochastic independent events sounds almost like a "guru's recommendation", and it is not useful to clarify the reason why. He concludes that the class needed to think about it, and he moves the discussion forward. The teacher could have accepted G1's arguments and praised it as the correct resolution. On the contrary, he stops the conversation, without commenting on the correctness of the resolution. The teacher's feedback is at the level of the process. He evaluates Enrico's reaction as being suitable for the situation and he considers Giacomo's intervention as not helpful to build common knowledge. Doing so, the teacher suggests that the function of a good argumentation is not only to support a resolution, but also to help students develop a common knowledge and understanding. The feedback is about the students' performance (where the students are in their learning trajectory) and, at the same time, it is useful to clarify the teaching and learning goals.

G2's representations

Later in the same lesson Enrico's group (G2) presents Fermat's resolution by means of a tree diagram represented in Figure 1. Fermat imagines that the players play the maximum number of games remaining and considers the possible outcomes. With a score of 5 to 3, there are three rounds left and eight possible outcomes, represented by the branches of the tree diagram at the top of Figure 1, where T is for heads and C is for tails. In seven out of eight cases A wins; that is, the probability of winning for A is 7/8 and for B is 1/8. Fermat's resolution is based on a fiction (in any case, the two players play three rounds) that is not easy for students to accept. In fact, G2-students observe that the players would play three more rounds only if tails came up in the first two tosses, otherwise the game stops after one or two rounds. Then, they present another possible representation of the game (Figure 1, lower part). In this tree diagram

each branch stops with the victory of one of the two players, and the probability of winning for B is calculated to be equal to 1/4.

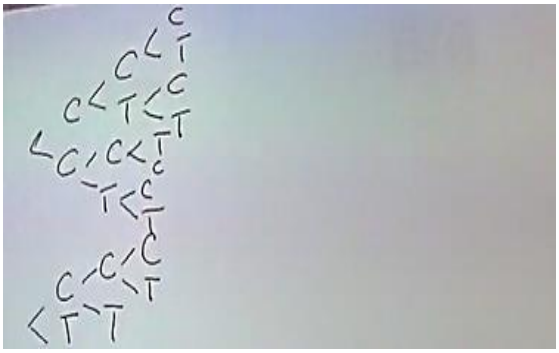


Figure 1: Tree diagrams of Fermat's (at the top) and G2's (below) resolutions.

G2's representations express their point of view and conviction about the problem, and their reluctance to accept Fermat's mathematical fiction. Their difficulty in grasping the not equiprobability in the second representation could be related to the *equiprobability bias* (Batanero et al., 2005). These representations are subsequently referred to by the teacher and peers and can be considered shared within the class.

Episode 2

The second episode takes place during step 4 of the activity. After small groups' reflection students are asked to discuss collectively. At a certain point of the class discussion, G1 shares their reflection and conclusion. Sabrina summarises the group resolution. She states that also using the G2's tree diagram it would be possible to calculate player B winning probability to be equal to 1/8 (and not 1/4). She supports her claim with the argument reported below. Her speech is intertwined with the construction of the pie chart illustrated in Figure 2, the steps for its construction have been included within the transcript in brackets.

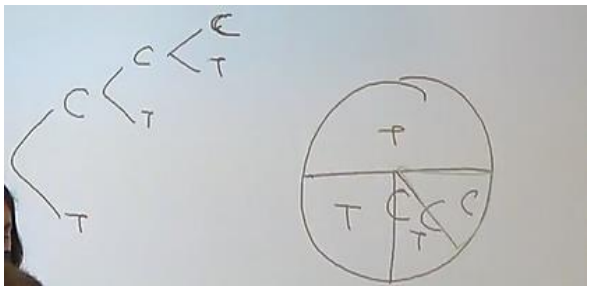


Figure 2: Group 1's pie chart.

Sabrina: We tried to explain it with a pie chart. Here there is a 50% chance to get heads or tails, so the chart is like this (she draws a circle and divides it in half with a segment. She writes T in the upper half of the circle and C in the lower one). If heads wins, the game ends here (she points at the upper half of the circle), but if tails wins there is again a 50% chance that heads wins or tails wins (she divides the bottom half in two equal parts and writes T for heads and C for tails). Again, here the game ends and here it continues. And then again here it splits because if heads come up, heads

> wins, or tails, tails wins (she divides the circle quarter denoted by letter C in two parts), either way the game ends, but what we see here is 1/8.

All the students in the classroom, particularly G2's students, claim they understand, and they agree with G1's solution. Sabrina's argument refers to G2's tree, which reflects their convictions and responds to their need to use representations more closely related to the game development. G1's representation allows Sabrina to show the not-equiprobability of the possible outcomes by correlating the branches with different-sized slices of the pie chart, and without referring to unshared theoretical results. Pie charts are generally used to represent ratios and quantities and students had no difficulty in reading the result in such representation. According to Mercier and Sperber (2017), it can be considered a good argumentation, since it shows the consistency between G1's and G2's convictions and knowledge, and it allows students to evaluate the relationships on their own.

DISCUSSION AND CONCLUSION

Andrea's and Sabrina's arguments support the same conclusion. However, they express their conclusion in a completely different way. If the function of argumentation is to help collective sensemaking development of probability theory concepts, it is crucial to consider peers' convictions and knowledge as a starting point. This allows to convince them by showing coherence relationships that they can evaluate on their own. Andrea's and Giacomo's arguments do not support the desired educational goal of developing probability concept sensemaking; therefore, they are rejected by some of their classmates and by the teacher. Enrico's reaction to Andrea's e Giacomo's interventions could be interpreted as a sign of the fact that Enrico shares the learning goals. Conversely, Sabrina's diagram seems to play a crucial role in conveying probability meanings. The pie chart strongly characterises Sabrina's argument and allows her classmates to change their opinion. The analysis shows that goals shared and expressed by the feedback seem to be crucial for shaping the development of the collective argumentation, clarifying what can be accepted or not by the classroom at the moment. Considering the functions that the argumentative texts should fulfill in the classroom could help the teacher give feedback to students. Moreover, since the development of argumentation depends on context and interlocutors, adapting one's arguments to different interlocutors could be understood as a feature of argumentative competence. In conclusion, we remark how interactions are "elegant ways to divide cognitive labor" (Mercier and Sperber, 2017, p. 236), and how this interactive nature of the process could help students refine and enhance the arguments produced. However, the characteristics of a suitable argument depend on several factors, such as the didactical functions of argumentative processes, and the time in the teaching-learning process at which they take place. Further investigations are certainly needed. Firstly, the role of representations and, in general, the possible ways in which coherence relationships can be shown in mathematical argumentation are to be explored further. Secondly, considerations about possible ways to enhance students'

participation as an attentive audience are to be made. Finally, the peculiarity of the probability context is to be deepened, and other contexts could be taken into account.

REFERENCES

Batanero, C., & Sanchez, E. (2005). What is the Nature of High School Students' Conceptions and Misconceptions About Probability?. In G. A. Jones (Ed.), *Exploring probability in school* (pp. 241–266). Springer, Boston, MA.

Borovcnik, M., & Kapadia, R. (2014). A historical and philosophical perspective on probability. In E. J. Chernoff & B. Sriraman (Eds.), *Probabilistic thinking* (pp. 7–34). Springer, Dordrecht.

Cusi, A., Morselli, F., & Sabena, C. (2017). Promoting formative assessment in a connected classroom environment: Design and implementation of digital resources. *ZDM*, 49, 755–767.

Ferrari, P. L. (2019). Argomentare a scuola: intrecci fra matematica e lingua. *L'insegnamento della matematica e delle scienze integrate, 42(5)*, 611–625.

Ferrari, P. L. (2021). *Educazione matematica, lingua e linguaggi*. Utet Università.

Hanna, G. (2020). Mathematical proof, argumentation, and reasoning. In S. Lerman (Ed.) *Encyclopedia of mathematics education* (pp. 561-566). Springer Cham.

Krummheuer, G. (1995). The ethnography of argumentation. In P. Cobb & H. Bauersfeld (Eds.), *The emergence of mathematical meaning: Interaction in classroom cultures* (pp. 229–269). Hillsdale, NJ, USA: Lawrence Erlbaum Associates.

Mariotti M. A. (2022). *Argomentare e dimostrare come problema didattico*. Utet Università.

Mercier, H., & Sperber, D. (2017). *The enigma of reason*. Harvard University Press.

Paola, D. (2019). Un'esperienza di avvio al pensiero probabilistico nella prospettiva di educare alla razionalità. In F. Morselli, G. Rosolini, & C. Toffalori (Eds.) *Educare alla razionalità: Tra Logica e Didattica della Matematica*. UMI.

Pedemonte, B. (2007). How can the relationship between argumentation and proof be analysed?. *Educational studies in mathematics, 66(1)*, 23–41.

Schwarz, B. B. (2009). Argumentation and learning. In N. M. Mirza, A. N. Perret-Clermont (Eds.), *Argumentation and education* (pp. 91–126). Springer New York.

Stylianides, A. L. (2007). Proof and proving in school mathematics. *Journal for research in Mathematics Education, 38*(3), 289–321.

Stylianides, A. J., Bieda, K. N., & Morselli, F. (2016). Proof and argumentation in mathematics education research. In Á. Gutiérrez, G. C. Leder, & P. Boero (Eds.), *The second handbook of research on the psychology of mathematics education* (pp. 315–351). SensePublishers Rotterdam.

Yackel, E. (2001). Explanation, Justification and argumentation in mathematics classrooms. In M. Van den Heuvel-Panhuizen (Ed.), *Proceedings of the 25th conference of the international group for the psychology of mathematics education PME-25*, (vol. 1, pp. 9–24). Utrecht University.

DIDACTIC-MATHEMATIC KNOWLEDGE TRAITS OF PRE-SERVICE TEACHERS WHEN POSING AND SOLVING ROBOTIC PROBLEMS

Gemma Sala-Sebastià[1], Adriana Breda[1], Alicia Sánchez[1], and Vicenç Font[1]

Universitat de Barcelona (Catalunya, Spain)

This paper aims to characterize traits of didactic-mathematical knowledge of pre-service teachers when solving and posing robotics tasks. We qualitatively identified aspects of mathematical and computational knowledge of 97 pre-service teachers of Early Childhood Education when solving tasks as users of the educative robot Bluebot, and we analyzed their justifications when reflecting on the design of robotic tasks. The results show that participants present characteristics of didactic-mathematical knowledge, but errors and ambiguities are evidenced in the programming procedures and representations. These errors and ambiguities influence the didactic suitability of the robotic tasks they designed. For further research, it is considered to develop didactic-mathematical and computational knowledge in the training of future teachers.

INTRODUCTION

Computational Thinking (CT) should be developed progressively, starting at an early age (Zapata-Ros, 2019). The current Spanish curriculum incorporates CT in the second cycle of Early Childhood Education (students from 3 to 6 years old), in the area of discovery and exploration of the environment, to develop in students the process of problem-solving (MEFP, 2022).

To attend to this normative demand, there is a need to introduce teacher training programs to foster CT from Early Childhood Education (among others, Benton et al., 2017; Ribeiro et al., 2011; Seckel et al., 2022a). In addition, research can be found that aimed to study pedagogical practices and teachers' conceptions regarding the use of robots in the early ages (among others, Papadakis, 2020; Seckel et al., 2021). These studies have highlighted the research agenda to study the knowledge that teachers should acquire to be able to teach. Within the framework of the Ontosemiotic Approach (OSA) (Godino et al., 2007), there is a model of Didactic-Mathematical Knowledge (DMK) that interprets and characterizes the teacher's knowledge (Pino-Fan & Godino, 2015). In that sense, the general objective of this research is to characterize traits of the didactic-mathematical knowledge of future kindergarten teachers when solving and posing robotic tasks.

THEORETICAL FRAMEWORK

The Didactic-Mathematical Knowledge (DMK) model interprets and characterizes the teacher's knowledge from three dimensions: mathematical dimension, didactic dimension, and meta-didactic-mathematical dimension (Pino-Fan & Godino, 2015). The mathematical dimension of the DMK includes common content knowledge and

2023. 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. 139-146). PME 46.

extended content knowledge. The didactic dimension of the DMK includes epistemic, cognitive, affective, interactional, mediational, and ecological subcategories of knowledge. The meta-didactic-mathematical dimension is the one that characterizes the knowledge that teachers need to reflect on their practice and to evaluate the didactic suitability to find possible improvements in the design and implementation of these processes (Breda et al., 2017).

For each of the components of the DMK, the OSA has "theoretical-methodological" tools that have been described and used in several investigations (Godino et al., 2007). For example, the "ontosemiotic configuration" tool is available for the development of instruments to systematically analyze teachers' knowledge of the mathematical dimension and the epistemic facet of the DMK. This tool allows the description and characterization of the primary mathematical objects-representations/language in its different registers; problem situations; concepts and definitions; propositions; procedures; and arguments/justifications-that are produced through mathematical processes or as part of the planning of a task (or sequence of tasks) for the classroom (Malaspina & Font, 2010). In addition to this, mathematical knowledge contemplates the description of teachers' errors and ambiguities from a mathematical point of view.

On one hand, for the development of instruments to systematically evaluate and analyze teachers' knowledge of the meta-didactic-mathematical dimension, the Didactic Suitability Criteria (DSC) tool becomes operational, which according to Font et al. (2010) are characterized as follows: Epistemic Suitability, to assess whether the mathematics being taught is "good mathematics"; Cognitive Suitability, to assess, before starting the instructional process, whether what is to be taught is at a reasonable distance from what the students know, and after the process, whether the learning is close to what was intended to be taught; Interactional Suitability, to assess whether the interactions resolve doubts and difficulties of the students; Mediational Suitability, to assess the adequacy of the material and temporal resources used in the instructional process; Affective Suitability, to assess the involvement (interests and motivations) of the students during the instructional process; Ecological Suitability, to assess the adequacy of the instructional process to the educative project of the center, the curricular guidelines, the conditions of the social and professional environment. The notion of DSC has had a relevant impact on teacher training in different countries, in particular, in the development of the teacher's meta-didactic-mathematical knowledge about their practice as Early Childhood pre-service teachers in Catalunya (Sala-Sebastià et al., 2022).

On the other hand, Estebanell et al. (2018), propose a model for initial teacher training that contemplates four levels for the development of CT: user, reflective user, teacher, and reflective teacher. Concerning the user level, Seckel et al. (2022b) propose a classification of the teachers' errors in the planning of programming, they are the following: a) error due to the absence of a function in programming; b) error by excessive quantification of a function in programming; c) error due to the misunderstanding of a type of programming, and d) error when applying previous

knowledge. Further, teachers must recognize two basic aspects when designing tasks (Arlegui & Pina, 2016). The first aspect is related to the idea of a robotics problem-type task, whose resolution involves the robot moving from an initial state to a final one, through the planning of a sequence of actions (intermediate states that are programmed). The second aspect is related to criteria that should guide the approach of a problem or a sequence of problems, these are: 1) to contemplate progressive complexity, 2) to refer to known and unknown aspects and 3) to place the problem in an environment (scenario).

METHODOLOGY

The participants in the study were 97 students of the Didactics of Mathematics course of the Early Childhood Education Degree at a Catalan university (Spain). The trainee teachers were organized into work teams to carry out the didactic sequence that was proposed to them; there were 17 work teams in total.

The data were obtained from the recordings of some of the sessions implemented, from the field notes of the first and second authors, who acted as teachers of the group carrying out participant observation, and from the documents written by the prospective teachers, called [D1], [D2] and [D3].

The learning sequence designed by the first two authors of the article and implemented with prospective teachers contemplated two sessions and autonomous work of the teams. The didactic objective of session 1 focused on the development of logical-mathematical and computational thinking skills in 5 or 6-year-old children by solving tasks with an educative robot. It was explained to them that they had to address and answer the questions in the dossier [D1] and describe what they had done to answer them, and no further information was given, neither about the working of the Blue-Bot robot nor about the use of the other resources provided. Each task in the dossier, in turn, responds to certain specific objectives of CT development that connect diverse contents in a transversal way, such as argumentation, logical-mathematical reasoning, and spatial and metric reasoning, among others. The purpose of the second session was the participants' identification of key aspects of the characteristics that a problem-type task should have and the sub-processes that should be carried out to solve it. Students were asked to think of characteristics of a good problem to be used for didactic purposes. Afterwards, each group should present a document with their list in a forum of the subject platform to share it with other participants [D2]. After this session, participants were asked to design as a team a session for 5 - 6-year-old children with educative robots that contemplates elements of the didactic use of the problem-type tasks identified in session 2 and the experience lived in session 1 [D3].

To identify aspects of mathematical knowledge presented by future kindergarten teachers when solving tasks as users of the Blue-Bot educative robot, the document [D1] with the answers of the future teachers to the tasks presented in the dossier and the explanation of the actions performed to achieve them were considered. Also, the recordings of the development and resolution of the tasks posed in the dossier were

considered when the future Early Childhood Education teachers manipulated the Blue-Bot robot. To identify traits of mathematical and computational knowledge, the notion of ontosemiotic configuration in the DMK model of the OSA (Pino-Fan & Godino, 2015) was used. In addition, the categories of errors that teachers make when programming the robot proposed by Seckel et al. (2022b) were used.

To collect all the characteristics that the participants believed that a problem-type task should have, a content analysis of this document [D2] was carried out, where we identified the emerging characteristics and which of them coincided among the different teams. Then, the design of a robotic task and its possible implementation with 5 - 6-year-old children made by the participants [D3] was taken into account. This document was analyzed to identify which of the emerging characteristics resulting from the analysis [D2] appeared in [D3], i.e., whether the pre-service teachers had considered the key aspects of the characteristics that a robotic problem should have. And then, it was analysed to identify the suitability of the participants' designs and to determine their mathematical meta-didactic knowledge.

Finally, a triangulation of the analyses among the most expert authors in the use of the instruments is carried out and it was possible to infer aspects of the didactic-mathematical and computational knowledge of the participants when posing and solving robotic tasks from the aspects of mathematical and mathematical meta-didactic knowledge emerging in the previous analyses. This inference was also triangulated with the opinion of an expert in the OSA theoretical framework.

RESULTS AND DISCUSSION

Firstly, we found arguments/justifications, representations/language, propositions and procedures referring to the mathematical objects sequence and measurement from the primary objects that emerge from the future teachers' answers to the proposed robotic problems. It was not possible to identify, for example, definitions or concepts referring to a specific mathematical object.

Many groups do not reach the level necessary to argue and justify whether different orders can take the robot to the same place. Although some teams already establish a justification based on the sequencing related to CT, many teams argue that the robot arrives at a certain place based on the concrete idea of a path (not referring to the robot's programming). Regarding the procedures pointed out by the teams for the calculation of the distance between one class and another, although the estimation procedure and the arithmetic calculation of multiplication appear, the idea of measurement by comparison (once they work with medium-scale distances) does not arise, for example.

In addition, most of the teams showed errors in the programming procedure when they had to give the instructions for the robot to return along the outward route. Regarding how the participants represent the orders given to the Blue-Bot, although the verbal-written, symbolic, and iconic representations appear, they present errors in the representation of the route that the robot follows and also ambiguities in the drawing of the arrows made with their iconic representation.

Although studies show positive levels of interest, knowledge, problem-solving and self-efficacy of future teachers when working with robotics and CT in basic education (Piedade, 2021), this study shows a lack of mathematical (sequence and measurement) and computational (sequence programming) knowledge when solving robotic problems at the user and reflective user level. This is consistent with the studies of Seckel et al. (2022b), who observe different types of errors made by future kindergarten teachers when solving robotic problems, both in the use of robot commands and their respective programming representation.

In addition to the results referring to the analysis of the mathematical and computational knowledge of the participants that show errors and ambiguities in the programming procedures, their designs lack didactic objectives focused on the teaching and learning of computational knowledge. The representation or writing of robot programming algorithms is not promoted as an essential element for the development of computational thinking. It can be inferred that future teachers do not contemplate it in their designs either because they do not master it (as they make mistakes), or because they are not aware of the importance of fostering the writing or representation of algorithms as an institutionalization of CT knowledge. This is reflected in the epistemic and cognitive suitability of the designs.

The designs take special care of aspects related to affective suitability, including many elements to motivate and involve children in a fun activity that generates positive emotions. To this purpose, they mainly provide task contexts (pupils' centres of interest; very close, everyday and familiar situations; games with prizes; children's stories, etc.), which aim to involve pupils in the activity. The resources identified in the analysis of mediational suitability support this focus on the emotional part of the design, since, in general, most of the resources are used to enrich the context (carpets decorated with motifs of the topic of the centre of interest, children's stories, etc.). However, the programme cards that could have been proposed as facilitating resources for learning and developing CT do not play a central role, as they are included only as an optional use. To get children involved in the activities designed, pupils are given an active and leading role in making decisions and expressing justifications and arguments. To do this, they are organized into small working groups that facilitate the necessary interaction (interactional suitability) for these situations of debate and reflection, led by the teacher, to take place. This type of design of future Early Childhood Education teachers with a strong concern about high affective suitability was also found in Sala-Sebastià et al. (2022).

According to the results, the least considered suitability by the pre-service teachers is the ecological suitability, as it seems that none of the participants consulted the curriculum to ensure that their proposal complied with the legal guidelines. Furthermore, although all designs include contextualized tasks, none of them shows an interdisciplinary approach. Only one deal with the environment in a cross-cutting manner but with little depth, and a couple of them foster connections with other content within mathematics (numerical and spatial thinking).

When future teachers design educational practices without having been given a detailed guide on how the designs should be, many of the elements of consensus in the educational community emerge, such as the fact that the task should involve and motivate students, that students should interact, share ideas and help each other, among others (Breda, et al., 2017). But other equally important aspects do not always emerge, such as, for example, establishing didactic objectives and assessment criteria, establishing a sequence of activities of progressive difficulty, establishing mechanisms for institutionalising the central contents that are the objective of educational practice, among others (Seckel et al., 2022).

CONCLUSIONS

The results show that future teachers present characteristics of didactic-mathematical and computational knowledge since, in the participants' answers, aspects of such knowledge can be timidly observed. However, they show certain weaknesses in mathematical and computational knowledge that are reflected in the results on the didactic suitability of the designs drawn up. So, we wondered whether it might be necessary to include in the training of these future teachers basic mathematical knowledge and, specifically, computational thinking to guarantee the quality of their future educative practices and to comply with the current guidelines of the Spanish and Catalan curricula.

If the DSC, which in this work has been used as a methodological element to carry out the data analysis, were made available to future teachers in the degree, they could use them as a design tool and implementation guide to help keep in mind and balance the various didactic facets of educational practice. Thus, the development of didactic knowledge in future teachers would be promoted.

The results of the study, determined from a particular context, have limitations as they are based on future teachers of kindergarten in a specific geographical area of Spain. Different results could be obtained if the study were carried out with future teachers of primary school or with in-service teachers in another location. It is also important to underline that the robotic tasks proposed in the dossier also conditioned certain types of response. Changing them could imply some modifications, albeit subtle, in the results found.

Further research could consider that the future teachers participating in the study carry out a kind of simulation by applying the tasks designed with their peers to bring out characteristics of didactic knowledge that were not contemplated in the design. In addition, it is considered relevant to incorporate mathematical and computational knowledge in the training of future teachers to operate the Blue-Bot robot and develop logical, spatial and computational thinking. It is also considered to incorporate the DSC in the training to develop didactic knowledge.

ACKNOWLEDGEMENTS

This research was funded by the Spanish R&D project PID2021-127104NB-I00 (MINECO/FEDER, EU).

REFERENCES

Arlegui, J., & Pina A. (2016). *Didáctica de la robótica educativa: Un enfoque constructivista*. Dextra Editorial S.L.

Benton, L., Hoyles, C., Kalas, I., & Noss, R. (2017). Bridging primary programming and mathematics: Some findings of design research in England. *Digital Experiences in Mathematics Education*, *3*(2), 115-138. https://doi.org/10.1007/S40751-017-0028 -X

Breda, A., Pino-Fan, L. R., & Font, V. (2017). Meta didactic-mathematical knowledge of teachers: criteria for the reflection and assessment on teaching practice. *EURASIA Journal of Mathematics, Science and Technology Education*, *13*(6), 1893-1918.

Estebanell, M., López, V., Peracaula, M., Simarro, C., Cornellà, P., Couso, D., González, J., Alsina, A., Badillo, E., & Heras, R. (2018). Pensament computacional en la formació de mestres. In *Guia didàctica*. Servei de Publicacions UdG.

Font, V., Planas, N., & Godino, J. D. (2010). Modelo para el análisis didáctico en educación matemática. *Infancia y Aprendizaje*, *33*(1), 89-105. https://doi.org/10.1174/021037010790317243

Godino, J. D., Batanero, C., & Font, V. (2007). The onto-semiotic approach to research in mathematics education. *ZDM–International Journal on Mathematics Education*, *39*(1-2), 127-135. https://doi.org/10.1007/s11858-006-0004-1

Malaspina, U. & Font, V. (2010). The role of intuition in the solving of optimization problems. *Educational Studies in Mathematics*, *75*(1), 107–130.

Ministerio de Educación y Formación Profesional (MEFP). Real Decreto 95/2022, de 1 de febrero, por el que se establece la ordenación y las enseñanzas mínimas de la Educación Infantil, 2022. https://www.boe.es/eli/es/rd/2022/02/01/95

Papadakis, S. (2020). Robots and Robotics Kits for Early Childhood and First School Age. *International Association of Online Engineering*. https://www.learntechlib.org/p/218338/.

Piedade, J. M. N. (2021). Pre-service and in-service teachers' interest, knowledge, and self-confidence in using educational robotics in learning activities. *Educação & Formação*, *6*(1), e3345. https://doi.org/10.25053/redufor.v6i1.3345

Pino-Fan, L. R., & Godino, J. D. (2015). Perspectiva ampliada del conocimiento didáctico-matemático del profesor. *Paradigma*, *36*(1), 87-109.

Ribeiro, C. R., Coutinho, C. P., & Costa, M. F. M. (2011). A robótica educativa como ferramenta pedagógica na resolução de problemas de matemática no ensino básico. In *Proceedings of the 6th Iberian Conference on Information Systems and Technologies* (pp. 15-18).

Sala-Sebastià, G., Breda, A., & Farsani, D. (2022). Future early childhood teachers designing problem-solving activities. *Journal on Mathematics Education*, *13*(2), 239-256.

Seckel, M. J., Breda, A., Font, V., & Vásquez, C. (2021). Primary school teachers' conceptions about the use of robotics in mathematics. *Mathematics, 9*(24), 3186.

Seckel, M. J., Breda, A., Farsani, D., & Parra, J. (2022a). Reflections of future kindergarten teachers on the design of a mathematical instruction process didactic sequences with the use of robots. *EURASIA Journal of Mathematics, Science and Technology Education, 18*(10), em2163.

Seckel, M. J., Vásquez, C., Samuel, M., & Breda, A. (2022b). Errors in programming and ownership of the robot concept made by trainee kindergarten teachers during induction training. *Education and Information Technologies, 27*(3), 2955-2975. https://doi.org/10.1007/s10639-021-10708-8

Zapata-Ros, M. (2019). Computational Thinking Unplugged. *Education in the Knowledge Society, 20*, 18–29. https://doi.org/10.14201/eks2019_20_a18

INDIVIDUAL CONCEPTION FRAMES AS A CONCEPT FOR THE ANALYSIS OF MATHEMATICAL LEARNING

Alexander Salle* and Marcus Schütte**

*University of Osnabrück (Germany); **University of Hamburg (Germany)

The mathematical thinking and acting of learners are determined by individual interpretations of mathematical concepts. These interpretations can be reconstructed from an individual perspective—as individual conceptions—as well as from an interactionist perspective—as specific frames. So far, however, it remains unclear, how individual conceptions take effect in mathematical interactions and, on the other hand, how joint negotiation processes about mathematics affect the development of sustainable individual conceptions. Combining both concepts, we develop the concept of individual conception frames in order to create a theoretical foundation for the empirical analysis of individual mathematical learning in interactions.

1. BASIC IDEAS AND INDIVIDUAL CONCEPTIONS OF MATHEMATICAL CONTENT

Basic ideas of mathematical concepts can be understood as interpretations of mathematical concepts which are ideal-typical. These interpretations are connected to real-life as well as to mathematical contexts and representations and can thus give meaning to the mathematical concept. Through these connections, basic ideas should also enable flexible mental variations of the concept at hand and their application in inner-mathematical as well as modelling contexts. (Blum & vom Hofe, 2016; vom Hofe 1995; Greefrath et al., 2021). Examples of basic ideas for subtraction of natural numbers are i) the interpretation of subtraction as the determination of a difference between two quantities, and ii) the interpretation of subtraction as the removal of a subset from an initial quantity. Basic ideas can thus answer, in particular, the question of which interpretations of the concept at hand learners should develop when dealing with mathematical contents.

The derivation and formulation of basic ideas are carried out through a factual analysis of the mathematical content: By relating definitions of the mathematical concept to real-life or mathematics-related contexts and phenomena (Freudenthal, 1983), researchers derive relevant interpretations of the concept and classify them based on commonalities of the respective interpretations (Salle & Clüver, 2021; vom Hofe & Blum, 2016).

However, basic ideas usually do not go along with what learners develop for actual interpretations of mathematical concepts. Therefore, in the relevant literature, the basic ideas are contrasted with *individual conceptions of mathematical concept* (vom Hofe, 1995; vom Hofe & Blum, 2016). Several empirical studies show the significance of such conceptions for the explanation of differences in students' performance (e. g.,

2023. 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. 147-154). PME 46.

Prediger, 2008; see also vom Hofe & Blum, 2016). Individual conceptions are understood as individual interpretations of one person which are also connected to real-life or mathematical contexts and representations. Individual conceptions are action-guiding for the individual (Balacheff & Gaudin, 2009; Jetses & Salle, submitted; vom Hofe, 1995). Regarding features of individual conceptions, two views can be identified in literature: The first view assumes that individual conceptions are to be regarded as situation-transcending constructs which are activated in different situations; the second view, which the authors of this paper follow, understands individual conceptions as situation-specific: individual conceptions are thus always newly constructed by a person in every situation and are therefore never the same (Jetses & Salle, submitted; Balacheff & Gaudin, 2009). Such situation-specific and action-guiding individual conceptions can be sustainable in terms of content, but also fragmented or flawed.

If, in similar situations, the same person shows similarities in the respective individual conceptions with regard to the underlying interpretation, we can speak of a common *core of individual conceptions* (Jetses & Salle, submitted). This core in turn has a character that transcends situations. Such cores provide information for researchers about which previous experiences and which previous knowledge could be guiding for the construction of corresponding individual conceptions of the person. The construction, further development and networking of sustainable individual conceptions is regarded as crucial for mathematical learning (Piaget, 1976; vom Hofe, 1995). Against this background, a crucial question is how learners can develop such cores of sustainable individual conceptions when learning mathematics at school in order to construct adequate individual conceptions on their basis.

Answers to this question can be given from an *individual perspective*, for example, through appropriately designed learning arrangements that focus in a special way on the construction of individual conceptions, such as functional thinking exercises that require mental changes of mathematical objects from geometry or other domains (e. g. Weber, 2007). However, the action-guiding character of individual conceptions is not limited to situations in which a person thinks about mathematics alone; in this sense, individual conceptions also influence interaction, i.e. the negotiation processes in exchange with teachers and other learners. That is why the question of how individuals construct and apply sustainable individual conceptions must also be answered from an *interactionist perspective* in particular. According to an interactionist understanding, individual interpretations of the corresponding mathematical concepts are influenced by the reciprocally referring actions of the interaction partners in interactions. This will also have an effect on the construction of individual conceptions or the core of precisely these of the participants. In contrast to other approaches that can be used to investigate the role of conceptions in interactions (e. g. Balacheff & Gaudin, 2009; Vergnaud, 1996), the concept of basic ideas and individual conceptions focuses on specific mathematical concepts. On the one hand, this allows a targeted view on these concepts, and on the other hand, enables comparisons of individual conceptions with the corresponding basic ideas.

The purpose of this paper is to formulate a theoretical foundation for a methodological approach which allows us to explore answers from an interactionist perspective by investigating cores of individual conceptions of specific mathematical concepts in interactions. This foundation shall take into account individual as well as interactionist aspects of the development of individual conceptions.

2. MATHEMATICS LEARNING FROM AN INTERACTIONIST PERSPECTIVE – THE CONSTRUCTION AND MODIFICATION OF FRAMES

At the center of an interactionist perspective on mathematics learning is the concept of learning mathematics in and by interaction. Implementations of an interactionist perspective on mathematics learning can be found in interactionist approaches of interpretive research in mathematics education (cf. Bauersfeld et al., 1988; Jung, 2019; Jung & Schütte, 2018; Jung et al., 2022; Krummheuer, 2007, 2011; Schütte, 2014; Schütte et al., 2021). Social interaction is following these approaches seen as the constituting starting point of mathematical learning processes. Essential theoretical assumptions of an interactionist perspective on mathematics learning can be found in the sociological theory of symbolic interactionism (Blumer, 1969, 2013) and in approaches to collective learning (Miller, 1986, 2006). According to Blumer, the coexistence of people is constituted by symbolic interactions, so that actions of the individual are always carried out on the basis of previously interpreted actions of the counterpart. In addition to the concept of interaction, the concept of interpretation and the meaning resulting from these interpretations thus form the foundations of symbolic interactionism. According to this theory, meanings of mathematical objects or operations can be described as "social products", which are produced in mutual collective processes of interpretation (Blumer, 2013, p. 67).

By participating in such collective constructions of meaning, individuals can transcend their limited individual abilities to construct meaning and in this way create the basis for individual mathematical learning (Schütte et al., 2021; Jung et al., 2022). However, from an interactionist perspective, learners do not enter collective negotiation processes without experience and knowledge, nor do they only build up situation-specific concepts in these processes. Therefore, at the beginning of an interaction, all learners create initial situation-specific interpretations of the situation they are in, so-called *definitions of situation*, based on their individual experiences and knowledge (Jung, 2019; Krummheuer, 1992). The learners involved in the negotiation processes always draft the definitions of situation in anticipation of possible interpretations of the other learners and change or adapt them in processes of mutual negotiation of mathematical meaning (Jung, 2019; Krummheuer, 1992) with other learners. Within this process of mutual alignment of individual definitions of situation, a moment of 'interpretive agreement', i.e. a shared interpretation between the learners—*an interpretive interim*, as Schütte et al. (2021, see also Jung et al., 2022) call it—can be achieved (on the term see also Voigt, 1994, p. 78; as *taken-as-shared* understanding). In the sense of Miller (1986, 2006), the interpretive interim holds the potential to be potentially innovative for the individual and thus to systematically transcend the

individual's interpretive capacities. In the sense of an interactionist understanding of learning, thus it constitutes the 'stimulation potential' for fundamental, situation-transcending cognitive construction and restructuring processes of the individual. Therefore, a one-time collective construction of an interpretive interim usually does not lead to individual cognitive construction or restructuring processes; rather, it requires the repeated negotiation of the same interpretive interim - not necessarily in the same group of learners. By that, the underlying situation definitions of the individual that led to the production of the interpretive interim can "converge" into more standardized and routinised retrievable situation definitions (Schütte et al., 2021; Jung et al., 2022). These permanent cognitive constructions or restructurings of learners are called the construction or modification of *frames* with reference to Goffman (Goffman, 1959; Krummheuer, 1992). On basis of such frames, learners remake their interpretations of the situation in subsequent mathematical negotiation processes and thus receive changed opportunities for participation. According to an interactionist perspective on mathematics learning, individual situation-transcending cognitive construction and restructuring processes are thus reciprocally connected with collective mathematical negotiation processes or structurally embedded in them (see also the concept of the commognitive framework by Sfard, 2008; Schütte et al., 2021).

3. JUXTAPOSITION OF INDIVIDUAL CONCEPTIONS AND FRAMES

If we compare the terminologies as well as the concepts for describing mathematics learning mentioned above, many similarities can be found in both, which are most notably reflected in the central role of interpretations for individual conceptions and frames. Moreover, 'blind spots' can be identified in both concepts, which could be illuminated by a possible combination. In this way, the concept of basic ideas makes it possible, to compare individual interpretations which are descriptively obtained with 'target' basic ideas which are derived normatively from mathematical considerations. Thus, the extent to which there is a certain match can be identified. So far, however, it has not been possible within the basic ideas framework to describe how learners will interpret the mathematical concept at hand in interactions in future, based on analysed cores of individual conceptions; the core of individual conceptions describes commonalities of different reconstructed individual conceptions only retrospectively so far. The framework of basic ideas cannot fundamentally describe how in interaction processes taking place within a group of learners—who contribute their own interpretations of mathematical concepts—affect and change the respective cores of individual conceptions of all learners. It can be assumed that cores of individual conceptions also become relevant for conceptions constructed in the future, for example, that learners use their previous experiences and knowledge, which have coagulated in cores, as a background for the construction of new conceptions—and this also and especially through exchange with others. However, a corresponding theoretical and methodological foundation for the empirical analysis of this connection is not available and will be developed in this paper in the following.

However, the analysis of situation-transcending cognitive construction and restructuring processes is precisely at the centre of the theoretical approach of mathematics learning in and by interaction, and in particular, of the central concept of frames. This could be a useful supplement to basic ideas: Thus, the theoretical approach of mathematics learning in and by interaction focuses the origin of the permanently reconstructed and modified individual restructuring the interaction with others. That is based on an understanding that without this interaction any frames would not be produced. Also, the concept of frame seems to allow describing how learners will interpret in interaction processes in future under the assumption that they will apply a reconstructed frame. The concept of frame conceptualises just a kind of background foil against which learners make future interpretations of situations, e.g. in further interactions. However, the concept of frame has a 'blind spot' in describing mathematics learning, too. According to the logic of the construction and modification of frames, these frames always receive a kind of collective validation of the quality of the emerging individual restructurings thorough the culture of involved mathematics practitioners. However, comparisons of learners' frames in terms of their quality with "target frames" that are derived normatively from mathematics, are not possible, although this would be desirable to assess mathematical learning potentials.

4. INDIVIDUAL CONCEPTION FRAMES

The considerations above thus give reason to expect that the merging of the concepts of frame and of individual conceptions core can fix the mutual blind spots in the reconstruction of mathematical learning processes. The theoretical merging builds on two parallels of the concepts. On the one hand, irrespective of their different fundamental theoretical foundations in interactionist or cognitive-constructivist approaches, both frames and cores of individual conceptions refer to the individual and make it possible to describe what is situation-transcending. On the other hand, both concepts have the basic assumption that they stem from a kind of bundle of interpretations: First, cores of individual conceptions can be understood as commonalities of a bundle of different individual interpretations of a mathematical concept (and are thus always related to this mathematical concept); secondly, frames can be understood as bundles of interpretive interims that an individual has produced in different situations with other learners involved in mathematical interactions (and thus refer to situations).

Based on these considerations, we introduce the concept of a frame that is shaped by a core of individual conceptions, or, in short, *individual conception frame* of a mathematical concept. This term attempts to address the 'blind spots' of both concepts. In the sense of an interactionist view of learning in relation to frames, this is an individual construct which is produced interactively. By that, it leads to expanded opportunities of participation in future interactions, in the way that the individual can build up on previously successful action-guiding interpretations of a specific mathematical concept. In the end, it has an effect on the interactive construction and modification of new individual frames of conception (Schütte et al., 2021). In the sense

of describing a core of individual conceptions that refers to a mathematical concept, an individual conception frame makes it possible to compare the individual construct which is interactively produced with normatively derivable basic ideas to this mathematical concept. In this way, making assessments about the sustainability of developed individual constructs of the learners becomes possible. Within this new conceptual network, basic ideas can be understood as a 'desirable' background foil that enables learners to construct ideal-typical interpretations of a specific mathematical concept in respective situations. Therefore, when comparing the individual conception frame with normative basic ideas, one can also speak of a comparison with a *basic idea frame* of the mathematical concept.

5. PERSPECTIVES

The purpose of this paper is to create a theoretical basis, on which mathematics learning from an individual and interactionist perspective can be described. With such a basis, new theoretical elements can be developed at the interface between theory and empiricism. Basic so-called sensitising theoretical concepts (Kelle, 1994, p. 239; cf. sensitising concepts Blumer, 1954, p. 7; Strauss & Corbin, 1994), which—like the individual conceptual frame here—are derived from existing theory, represent an important tool in this process. The basic assumption is that empiricism is able to disrupt and change these theoretical constructs and that they can be further developed into an empirically substantial concept through application in the empirical field.

In principle, on the one hand, we assume that learning mathematics through the construction of frames is a fundamentally interactively conditioned process. However, on the other hand, having regard to the concept of basic ideas, individual conceptions are also crucial for such learning. According to this understanding, any differences that arise from different frames—different individual conception frames in particular—of the individuals who are participating in the interaction are a motor of learning (Schütte, 2014). Such differences challenge new interpretative agreements of the participants and thus lead to the construction and modification of existing frames. How this construction and modification of individual conception frames can be described, how such a description can enable the reconstruction of mathematical learning at the interface between the individual and the interaction partners, and in how far eventual modifications of frames also become apparent in non-interactively shaped situations in which the mathematical concepts are applied, shall be subject of future empirical studies.For this purpose methodological approaches of both underlying research paradigms are to be brought into synergy with each other in the sense of a mixed methods design (Jetses et al., submitted; Schütte et al., 2019).

REFERENCES

Balacheff, N., & Gaudin, N. (2009). Modeling students' conceptions: The case of function. *CBMS Issues in Mathematics Education*, 16, 207–234.

Bauersfeld, H., Krummheuer, G., & Voigt, J. (1988). Interactional theory of learning and teaching mathematics and related microethnographical studies. In H.-G. Steiner, & A. Vermandel (Eds.), *Foundations and Methodology of the Discipline Mathematics Education* (pp. 174–188). University of Antwerp.

Blumer, H. (1969). *Symbolic Interactionism. Perspective and Method*. Prentice-Hall.

Freudenthal, H. (1983). *Didactical phenomenology of mathematical structures*. Reidel. https://doi.org/10.1007/0-306-47235-x

Goffman, E. (1974). *Frame analysis. An essay on the organisation of experience.* Harvard University Press.

Greefrath G., Oldenburg R., Siller, H.-S., Ulm V., & Weigand, H.-G. (2021). Basic mental models of integrals: theoretical conception, development of a test instrument, and first results. *ZDM – Mathematics Education*, *53*(3), 649–661. https://doi.org/10.1007/s11858-020-01207-0

vom Hofe, R. (1995). *Grundvorstellungen mathematischer Inhalte.* Spektrum.

vom Hofe, R., & Blum, W. (2016). "Grundvorstellungen" as a category of subject-matter didactics. *Journal für Mathematik-Didaktik, 37*(S1), 225–254. https://doi.org/10.1007/s13138-016-0107-3

Jetses, T. & Salle, A. (submitted). Rekonstruktion individueller Vorstellungen im Rahmen des Grundvorstellungskonzeptes. *Journal für Mathematik-Didaktik.*

Jetses, T., Salle, A., & Mikoleit, M. (submitted). Zur Anwendung fachlich korrekter Deutungen des Ableitungs- und Integralbegriffs. *Mathematica Didactica.*

Jung, J. (2019). Möglichkeiten des gemeinsamen Lernens im inklusiven Mathematikunterricht – Eine interaktionistische Perspektive. In B. Brandt & K. Tiedemann (Eds.), *Mathematiklernen aus interpretativer Perspektive I – Aktuelle Themen, Arbeiten und Fragen* (pp. 103–126). Waxmann.

Jung, J., & Schütte, M. (2018). An interactionist perspective on mathematics learning: conditions of learning opportunities in mixed-ability groups within linguistic negotiation processes. *ZDM – Mathematics Education, 50*(6), 1098–1099. https://doi.org/10.1007/s11858-018-0999-0

Jung, J, Schütte, M. & Krummheuer, G. (2021). Discourses as the Place for the Development of Mathematical Thinking – an interactionist perspective. In J. Ingram, M. Schütte & M. Ní Ríordáin (Eds.), *Proceedings of Topic Study Group 39* (pp. 61–68). 14th International Congress on Mathematical Education.

Kelle, U. (1994). *Empirisch begründete Theoriebildung. Zur Logik und Methodologie interpretativer Sozialforschung.* Deutscher Studien Verlag.

Krummheuer, G. (2007). Argumentation and participation in the primary mathematics classroom. Two episodes and related theoretical abductions. *Journal of Mathematical Behavior, 26,* 60–82. https://doi.org/10.1016/j.jmathb.2007.02.001

Krummheuer, G. (2011). Representation of the notion "learning-as-participation" in everyday situations of mathematics classes. *ZDM, 43,* 81–90. https://doi.org/10.1007/s11858-010-0294-1

Piaget, J. (1976). Piaget's Theory. In B. Inhelder & H.H. Chipman (Eds.), *Piaget and His School* (pp. 11–23). Springer. https://doi.org/10.1007/978-3-642-46323-5_2

Prediger, S. (2008). The relevance of didactic categories for analysing obstacles in conceptual change: revisiting the case of multiplication of fractions. *Learning and Instruction, 18*(1), 3–17. https://doi.org/10.1016/j.learninstruc.2006.08.001

Salle, A., & Clüver, T. (2021). Herleitung von Grundvorstellungen als normative Leitlinien – Beschreibung eines theoriebasierten Verfahrensrahmen. *Journal für Mathematik-Didaktik, 42*(2), 553–580. https://doi.org/10.1007/s13138-021-00184-5

Schütte, M. (2014). Language-related specialised learning in mathematics. A comparison of learning settings: family, nursery and primary school. *The International Journal on Mathematics Education (ZDM), 46*(6), 923–938. https://doi.org/10.1007/s11858-014-0632-9

Schütte, M., Friesen, R.-A., & Jung, J. (2019). Interactional analysis. A method for analysing mathematical learning processes in interactions. In G. Kaiser & N. Presmeg (Eds.), *Compendium for early career researchers in mathematics education* (pp. 101–129). Springer. https://doi.org/10.1007/978-3-030-15636-7_5

Sfard, A. (2008). *Thinking as communicating. Human development, the growth of discourses, and mathematizing*. Cambridge University Press. https://doi.org/10.1017/cbo9780511499944

Strauss, A. & Corbin, J. (1994). Grounded Theory Methodology: An Overview. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of Qualitative Research* (pp. 273-285). Sage.

Vergnaud, G. (1996). The Theory of Conceptual Fields. In L. Stette, P. Nesher, P. Cobb, G. A. Goldin, & B. Greer (Eds.), *Theories of Mathematical Learning* (pp. 219-240). Lawrence Eribaum Associates.

Voigt, J. (1995). Thematic patterns of interaction and sociomathematical norms. In P. Cobb, & H. Bauersfeld (Eds.), *The emergence of mathematical meaning: interaction in classroom cultures* (pp. 163-201). Hillsdale, N. J.: Lawrence Erlbaum.

Weber, C. (2007). *Mathematische Vorstellungen bilden. Praxis und Theorie von Vorstellungsübungen im Mathematikunterricht der Sekundarstufe II.* h.e.p. https://doi.org/10.1007/bf03339054

DIGITAL MONITORING OF FRACTION LEARNING: ADAPTING A TEST FOR KNOWLEDGE OF FRACTION SUBCONSTRUCTS

Constanze Schadl, Anke Lindmeier

University of Jena, Germany

Fraction learning is a central yet difficult topic in mathematics education. Hence, formatively monitoring fraction learning is important. Despite a sound research base from paper-based studies, findings on fraction learning are still difficult to use in practice, as mathematics teachers have limited access to test materials, for example. This contribution presents the first results of a study that explores the feasibility of digital testing with n = 231 sixth graders. It focuses on the psychometric quality of fraction tests that have been adapted to a digital format and their suitability for repeated testing. The findings indicate similarities but also differences compared with the paper-based tests. The two tests indicate to be parallel and hence suitable for the digital monitoring. The implications of our findings for school practice are discussed.

INTRODUCTION

Even though fraction learning is a central topic in mathematics education and predicts later learning success (Siegler et al., 2012), most students encounter a range of problems while learning fractions (e.g., McMullen & Van Hoof, 2020). Prior studies inform about the components of fraction knowledge and central mathematical predictors for successful learning in this area (e.g., Schadl, 2020). Whereas most research focusing on predictors is restricted to statements such as "*more* (prior knowledge) *is better* (for later fraction learning)", the recent literature provides models that allow describing the relations between central predictors and fraction knowledge based on levels models that capture qualitative differences in knowledge beyond that linearity (Schadl, 2020). Even though promising for the educational context, teachers cannot benefit from these findings in practice so far. For instance, access to test materials is often restricted, and paper-based tests are inconvenient to administer and evaluate. Providing tests for digital testing could be an opportunity to remedy these issues. Further, digital testing may allow the assessment of learning trajectories and supports mathematics teachers to formatively assess and monitor students' fraction learning (Black & William, 2004).

Following the three-staged development process of instruments for monitoring learning (Fuchs, 2004), research should first focus on the psychometric quality of developed tests (stage 1) in terms of validity and reliability, for instance, with item response theory (IRT) scaling methods. Next, it is required to investigate their suitability for monitoring learning progressions based on repeated testing (stage 2) and their usability by mathematics teachers (stage 3). In this contribution, we address stage 1 and touch stage 2, presenting preliminary analyses of the piloting of a digitized fraction test with two test forms (a short and a long one).

2023. 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. 155-162). PME 46.

Structural modeling of fraction knowledge

Several approaches can be found in the literature to describe the structure of fraction knowledge from different perspectives. So, for example, Siegler and colleagues (2011) emphasize students' understanding of fraction magnitude representations and fractions placing on number lines as central (magnitude perspective). Other approaches highlight the meaning of fractions (part-whole perspective), enabling students to gain deepened insights into fraction procedures. From this latter perspective, Schadl (2020) provides a three-facet model with *knowledge of fraction subconstructs*, *fraction arithmetic skills*, and *fraction word problems*. Developing fraction knowledge in this perspective depends on mathematical predictors such as whole number multiplication and division. This contribution follows the part-whole perspective and focuses on the knowledge of fraction subconstructs according to Schadl (2020).

Modeling knowledge of fraction subconstructs from low to high based on results from paper-based studies

Based on a paper-based large-scale assessment with $N = 782$ sixth and seventh graders, Schadl (2020) modeled knowledge of fraction subconstructs from low to high with four levels. The test materials used in this large-scale assessment contained tasks for the subconstructs part-whole, ratio, operator, quotient, and measure that were strongly, slightly, or non-prestructured. Strongly and slightly prestructured tasks required shading fractions on models that were divided into *equal* parts or to evaluate statements as (in)correct (both strongly prestructured task formats) or to select the correct statements from several given (slightly prestructured task format). Such pre-structures were not given in non-prestructured task formats, where students had to construct and present solutions. So, the original test spans different subconstructs of fraction knowledge and covers, in addition, items with varying degrees of demands for students.

The empirically supported hierarchical models suggest that level 1 is characterized by the part-whole subconstruct in strongly prestructured task formats. Level 2 is characterized by the part-whole subconstruct in slightly and non-prestructured task formats. Further, tasks that require linking the part-whole with other subconstructs in strongly prestructured situations are characteristic of level 2. On level 3, this linking is required either in slightly or non-prestructured task formats. On level 4, students have to be able to deal with complex relations between the subconstructs with no or slight prestructuring. To summarize, the levels differ in whether the tasks require a transfer between the part-whole and other subconstructs. Further, the degree of task prestructuring is a central criterion for level characterization in the paper-based setting.

THE PRESENT STUDY

This contribution focuses on a digitized test for knowledge of fraction subconstructs that was adapted from the paper-based test from Schadl (2020). Two parallel test forms, a long and a short version, were prepared for the intended use in the digital monitoring of fraction learning. In this study, we investigate questions of test quality related to the transfer of the test into a digital setting and its suitedness for repeated testing:

Question 1 (Q1, level modeling): Which levels can be characterized to model knowledge of fraction subconstructs from low to high based on the digitized tests? To what extent do these levels replicate the ones from the paper-based setting?

Question 2 (Q2, test parallelization): Are students assigned to similar performance levels by both tests indicating their suitability for digital learning monitoring?

METHOD

The systematic introduction of fractions is usually done in grade 6 in Germany and is not part of the curricula at the primary level. We conducted the study including three measurement points from June 2022 to July 2022 after the systematic introduction of fractions had been finished with students in grade 6. To answer our research questions regarding the suitability for monitoring learning, we repeatedly administered the tests within a time frame of about 3 weeks. So, we expected the students' knowledge to remain stable over this period. We used a test with 36 items (called long test) for the first and one with 21 parallelized items (called short tests) for the second and third measurement points. We used a whole-class setting, either in computer rooms or with tablets in classrooms and the sixth graders worked about 45 minutes on the long and about 25 on the short tests. Tasks were presented in a randomized order using the online platform Levumi (Mühling, 2019). This paper reports preliminary analyses for the long and short test forms used in the first and second points of measurement.

Sample

N = 231 sixth graders (48.5% female) from 9 German classes for higher education (Gymnasium) participated in the first and second points of measurement. 203 students (48.3% female) worked on both tests.

Instrument

The tests were adapted from the previous paper-based test for knowledge of fraction subconstructs (Schadl, 2020). So, the tests included items for the subconstructs part-whole, ratio, operator, quotient, and measure (see Table 1 for exemplary items). For each subconstruct, the items were strongly prestructured (e.g., subconstructs part-whole and operator in Table 1), slightly prestructured (e.g., subconstructs quotient and measure in Table 1), or non-prestructured (e.g., the subconstruct ratio in Table 1) as these tasks did not include any answer options. Further, we constructed different subtypes for each subconstruct (four subtypes for the part-whole and three for the other subconstructs each). Subtype 1 required judging statements as correct or wrong for all subconstructs (e.g., operator subconstruct in Table 1) except for the measure subconstruct. Regarding the measure subconstruct, subtype 1 required locating numbers (1 or fractions) on number lines. Subtype 2 required selecting the correct statements from several options for all subconstructs (e.g., quotient and measure subconstructs in Table 1) except for the part-whole subconstruct for which subtype 2 required determining shares in given models. For subtype 3, we used different item

formats for each subconstruct. For example, subtype 3 required to color fractions (part-whole subconstruct in Table 1) or determine ratios (ratio subconstruct in Table 1). Subtype 4 required determining the whole for a given part. For test parallelization, we systematically varied the iconic models (e.g., subconstruct part-whole in Table 1), the numbers (e.g., subconstruct measure in Table 1 where the fractions were replaced in the short test; further examples would be the subconstructs operator and quotient in Table 1), or the situation presented in the text. For example, regarding the exemplary task for the ratio subconstruct in Table 1, we replaced the swimming pool situation with a theatre situation and made necessary adaptions (e.g., seats instead of lockers, tickets sold instead of closed lockers). Beyond these systematic variations, digitizing the paper-based test required some adaptions as we intended to minimize possible difficulties when students have to enter their solutions.

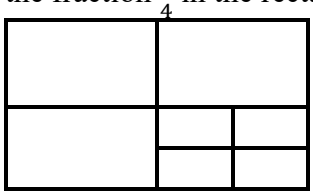
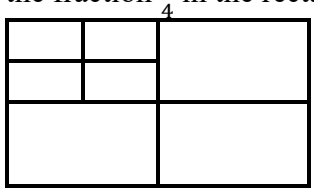
Sub-construct	Long test	Short test
Part-whole (10; 6)	Mark the fraction $\frac{1}{4}$ in the rectangle. 	Mark the fraction $\frac{1}{4}$ in the rectangle. 
Ratio (7; 4)		A swimming pool has 60 lockers. 40 lockers are closed at lunchtime. The ratio of open to closed lockers is the following: _____ : _____
Operator (6; 4)	If you multiply a number by five and then divide by six, you get the same result as if you divide the number by $\frac{5}{6}$. □ correct | □ wrong □ you cannot tell	
Quotient (6; 3)	$\frac{7}{10}$ is the result of □ 10 – 7 | □ 7 : 10 □ 10 : 7 | □ 7 – 10 □ all of the given calculations are wrong.	
Measure (7; 4)	How many fractions lie between the fractions $\frac{3}{5}$ and $\frac{4}{5}$? □ None | □ only a few | □ many □ infinitely many | □ none of the given answers is correct.	How many fractions lie between the fractions $\frac{7}{9}$ and $\frac{8}{9}$? □ None | □ only a few | □ many □ infinitely many | □ none of the given answers is correct.

Table 1: Sample tasks of the long and short test form. The numbers in brackets denote the number of items per subconstruct in the (long; short) test, with the long test consisting of 36 items and the short test of 21 items in total.

So, we replaced some open task formats with closed ones. For example, regarding the exemplary task for the measure subconstruct in Table 1, the task was presented without the answer options in the previous paper-based test. Typical wrong answers from the paper-based setting informed the answer options presented in the digital version.

Analysis

We evaluated the psychometric quality with the dichotomous Rasch model (Rasch, 1960) that allows describing with an additionally applied bookmark method (Mitzel et al., 2001) the knowledge of fraction subconstructs in levels. This report presents a common scaling for both tests and visualizes the results in a Wright Map. This map shows the persons on the left, with low-performing persons (corresponding to low person parameters) further down and high-performing persons (large person parameters) further up. Items are plotted on the right, with simple items (corresponding to low item parameters) plotted further down and difficult items (large item parameters) further up. If the person and item parameters are plotted at the same height, the corresponding persons are assumed to solve the items plotted at this height with a probability of 50%. If person parameters are larger than item parameters, the probability increases, and if person parameters are lower than item parameters, the probability decreases. Regarding the question of test parallelization, we used a graphical model test and plotted the person parameters of the short test against those of the long test. In addition, we calculated the correlation between the two test scores of the students.

RESULTS

We observed good fit parameters of the estimated IRT model with WLE-reliability of .81 and infits (weighted fits) ranging from 0.83 to 1.23. Regarding level modeling (see Figure 1), level 1 is characterized by the part-whole subconstruct *and* strongly prestructured task formats (subtypes 2 and 3). On level 2, tasks primarily require evaluating statements of fractions as correct or wrong (subtype 1). These statements refer to the subconstructs part-whole, ratio, and quotient.

On level 3, these statements also refer to the operator subconstruct or are more complex in that they refer to equivalent fractions, for example. Further, this level includes the tasks with the slightly prestructured format for the subconstructs ratio and quotient (subtype 2) that require selecting correct statements from several given, in the case of the quotient subconstruct also for equivalent fractions. Regarding the part-whole subconstruct, subtypes 2 and 3 are less prestructured than the tasks of level 1 as models are not divided into *equal* parts, for example (see subtype 3). Further, tasks require determining the whole of a given part (subtype 4). Regarding the measure subconstruct, tasks require locating the number 1 on number lines with different scalings.

Level 4 is characterized by subtype 2 for the subconstructs ratio, operator, and measure. Regarding the ratio subconstruct, tasks are more complex than in level 3, as ratios have to be used on level 4. Regarding the measure subconstruct, tasks also require locating fractions on number lines. Tasks of subtype 3 with non-prestructured task formats are

characteristic for level 4 regarding the quotient subconstruct and level 5 regarding the subconstructs ratio, operator, and measure.

In sum, the digitized levels are similar to those from the paper-based setting as the need to transfer between the part-whole and other subconstructs and the degree of prestructuring emerge as central for level characterization. The level models differ in the number of levels as we further subdivided the lowest level so that basic part-whole knowledge is indicative of the lowest level. Regarding the question of test parallelization (Q2), we observed similar mean overall scores in the long (M = 0.40, SD = 0.15) and short (M = 0.42, SD = 0.15) tests. Further, the bivariate scatterplot (see Figure 2) indicates students to be assigned to similar levels in both tests. This is also supported by a strong correlation between the long and short-test achievement of each person with $r(201) = .63$, $p < .01$.

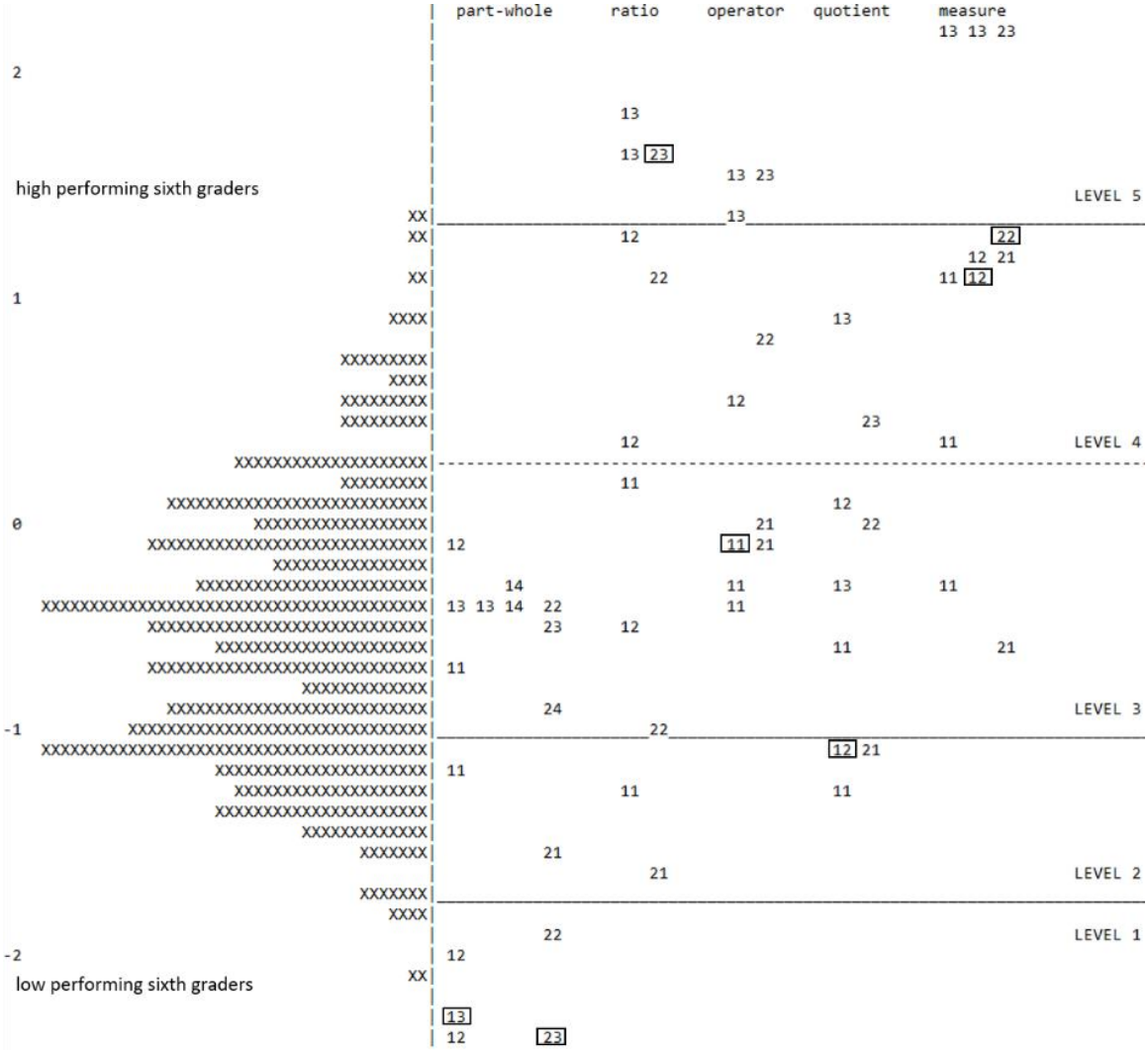


Figure 1: Wright-map of the IRT scaling results with items ordered according to the subconstructs. The first number indicates whether the item is part of the long (1) or the short test (2), and the second indicates the subtype. Sample items from Table 1 are boxed.

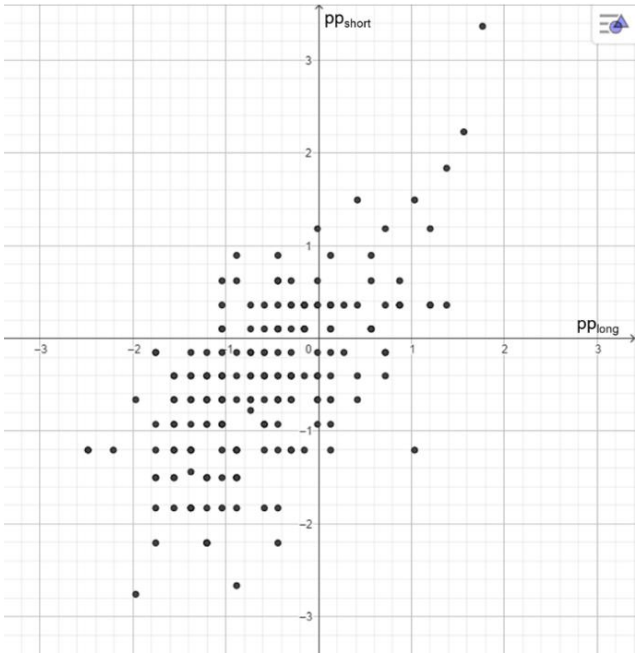


Figure 2: Bivariate scatterplot with the person parameters of the short test plotted against those of the long test.

DISCUSSION

The literature provides a range of relevant findings about fraction learning, primarily from paper-based studies. Digital technologies might be helpful to use the findings in practice to support students' learning, for instance, to make test materials accessible for mathematics teachers. Such access might support teachers in implementing formative assessment and allow the digital monitoring of students' fraction learning. The presented study is part of a larger research program to provide digital tests for use in practice. The report focuses on two parallel test forms, a long and a short version, for knowledge of fraction subconstructs that were adapted for use in a digital setting from a paper-based test. We investigated questions of test quality related to the transfer of the test into a digital setting and its suitedness for repeated testing. The adaption of the test for knowledge of fraction subconstructs to the digital setting was feasible. Despite the necessity to adapt test formats and include more parallel items, the analyses indicate that the test fits a model of knowledge of fraction subconstructs at different levels. We largely observed similarities of the resulting model to the model based on the previous paper-based setting, such as that it is central for level characterization whether the tasks required transferring between the part-whole and other subconstructs and whether the tasks were strongly, slightly, or non-prestructured. So, tasks of the part-whole subconstruct and strongly prestructured task formats seem to be easier than those that require linking the part-whole with other subconstructs, respectively slightly and non-prestructured formats. The quotient subconstruct emerges easier than other subconstructs, primarily those of operator and measure. Contrary to the paper-based setting, the data from the digitized tests suggests five instead of four levels with a finer resolution of basic abilities. Regarding repeated testing, students were assigned to similar levels based on the long and short tests in our study, which can be considered

first evidence that the theoretically parallel tests are indeed suited for digitally monitoring of fraction learning.

As the next steps, further analyses, including the second short test, are pending, and larger piloting studies, including students just learning fractions, are needed. Moreover, similar digital tests have to be developed for all facets of fraction knowledge, such as fraction arithmetic skills and fraction word problems, as well as their predictors. Even though the first research has been published in this context (Schadl & Lindmeier, 2022), the domain is still in its infancy. Investigations, like reported in this contribution, are necessary to ensure the suitability of the approach before the tests can be productively used in instruction and support students' learning. These first investigations, however, certainly show the potential that research findings can be used to build instruments for digital learning monitoring.

REFERENCES

Black, P., & William, D. (2004). The formative purpose: Assessment must first promote learning. *Yearbook of the National Society for the Study of Education, 103*(2), 20-50.

Fuchs, L. (2004). The past, present, and future of curriculum-based measurement research. *School Psychology Review, 33*(2), 188-192.

McMullen, J., & Van Hoof, J. (2020). The role of rational number density knowledge in mathematical development. *Learning and Instruction, 65*, 101228.

Mitzel, H., Lewis, D., Patz, R., & Green, D. (2001). The bookmark procedure: Psychological perspectives. In G. Cizek (Ed.), *Setting performance standards: Concepts, methods, and perspectives* (pp. 249-281). Lawrence Erlbaum Associates.

Mühling, A., Jungjohann, J., & Gebhardt, M. (2019). Progress monitoring in primary education using Levumi: A case study. In H. Lane, S. Zvacek, & J. Uhomoibhi (Eds.), *CSEDU 2019: Proceedings of the 11th International Conference on Computer Supported Education* (pp. 137-144). SCITEPRESS – Science and Technology Publications.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests.* The Danish Institute for Educational Research.

Schadl, C. (2020). *Individuelle Lernvoraussetzungen für den Erwerb des Bruchzahlkonzepts. Strukturanalysen und Untersuchung der längsschnittlichen Prädiktivität.* Waxmann.

Schadl, C., & Lindmeier, A. (2022). Modelling proportional reasoning skills in levels within a digital setting. In C. Fernández, S. Llinares, Á. Gutiérrez, & N. Planas (Eds.), *Proceedings of the 45th Conference of the International Group for the Psychology of Mathematics Education* (pp. 387-394). PME.

Siegler, R., Duncan, G., Davis-Kean, P., Duckworth, K., Claessens, A., Engel, M., Susperreguy, M., & Chen, M. (2012). Early predictors of high school mathematics achievement. *Psychological Science, 23*(7), 691-697.

Siegler, R., Thompson, C., & Schneider, M. (2011). An integrated theory of whole number and fractions development. *Cognitive Psychology, 62*, 273-296.

CHARACTERIZING EXTERNAL VISUALIZATION INTERVENTIONS: A SYSTEMATIC LITERATURE REVIEW

Johanna Schoenherr and Stanislaw Schukajlow

University of Münster, Germany

In this systematic literature review, we applied a qualitative content analysis to 13 research articles, comprising 17 external visualization (EV) interventions, published between 2018 and 2022. Our aim was to gain insights into the efficacy of EV interventions and to identify the EV intervention characteristics that might mediate the effect of the intervention on students' learning. We found that most EV interventions reported a positive impact on school students' understanding and problem solving. On the basis of our analyses and the explanations provided by the authors, we hypothesized nine characteristics of effective EV interventions, including the visualization process, technology use, multiple EVs, visual interaction, metacognitive reflection on EV, scaffolding of EV, and EV transfer.

INTRODUCTION

Mathematics as a discipline often deals with abstract objects and concepts, such as rational numbers or change. External visualization (EV; i.e., the use of physically embodied depictional representations (Schnotz, 2005), such as a tape diagram, a drawing, or a graph of a function) can be an important medium that supports students' learning (Arcavi, 2003). As two examples, exploring the relationships between the graph of a function and its derivative in an interactive learning environment can help students understand the concept of change, and translating a Bayes problem into the EV of a tree diagram might help them solve the problem. Consequently, mathematics instruction should provide students at all educational levels with learning environments that focus on EV to support their learning and to develop their abilities in constructing, using, and interpreting EV (OECD, 2019). In this systematic literature review, we analyze the extent to which recent mathematics education research has addressed EV interventions, and we synthesize the characteristics of the interventions and the studies' findings to gain insights into the efficacy of these learning environments.

THEORETICAL BACKGROUND

EV in mathematics education

In a previous scoping review (Schoenherr & Schukajlow, under review), we identified visualization components, tools, and purposes as three key characteristics of EV in recent mathematics education research: First, EV includes two visualization components (Arcavi, 2003). The process component includes all physical and mental activities and processes related to selecting, constructing, using, and interpreting EVs. The product component describes the resulting visual depiction (e.g., type, appearance, and accuracy of EV). EV interventions can focus on one or both components. For

2023. 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. 163-170). PME 46.

example, Kobiela and Lehrer (2019) had students experience and reflect on the process of generating a rectangle by asking students to sweep paint on a ceramic tile with a squeegee. Second, students made use of tools to interact with EVs, including paper-pencil, hands-on objects, gestures, or technology. EV interventions can give students opportunities to use one or several tools. Third, students use EV for diverse purposes (e.g., understanding, problem solving, and applying proofs in various mathematical content domains). Although EV seems to be most obvious in the domain of geometry, as geometry often relies on spatial reasoning, students can use EV in other domains, such as calculus and algebra (Arcavi, 2003). An open question is: To what extent does the efficacy of EV interventions depend on the intended purpose and content domain?

Learning environments that focus on EV

We define an EV intervention as any school, classroom, or learning environment in which students are provided with a set of activities aimed at promoting mathematics learning with or through the use of EV in an empirical study (e.g., a teaching sequence given to a class of Grade 12 students on the concept of change using an interactive graphing tool). Due to the variety of different kinds of EVs and different ways of implementing these interventions in learning settings, EV interventions can differ greatly. To the best of our knowledge, previous findings on the impact of EV interventions have not yet been synthesized. By systematically synthesizing recent intervention studies and their findings in the current review, we aim to describe learning environments that have focused on EV in recent mathematics education research and offer insights into their impact on student learning.

Characteristics of learning environments

Previous reviews in other research areas have revealed various mediating characteristics of powerful learning environments. For example, in a meta-analysis of 84 studies, Dignath and Büttner (2008) identified *duration of training* and *metacognitive reflection on learning* amongst others as characteristics of effective learning environments targeting self-regulated learning. As another example, Duijzer et al. (2019) compiled characteristics of embodied learning environments from 44 research articles, including the *real-world context*, *multimodality*, *multiple representations*, *student control*, and *attention capturing*. Besides these EV-unspecific characteristics of learning environments, little is known about the characteristics that are specific to EV for learning mathematics. As one example, Fiorella and Zhang (2018) discussed the *scaffolding of EV* and *metacognitive reflection of EV* as characteristics that potentially influence the efficacy of self-generated drawing for STEM learning. As another example, Presmeg (1986) concluded from an analysis of teaching styles that the generalization of specific EV is important for students to be able to learn with EV. An open question is: Which EV-specific and EV-unspecific intervention characteristics mediate the effects of recent EV interventions on student learning in mathematics?

RESEARCH QUESTIONS

We systematically reviewed recent empirical studies that were published in the last 5 years and investigated EV interventions that were designed to support school students' mathematics learning. Our aim was to address the following research questions: (a) What does the research literature on EV interventions report on their efficacy? (b) Which EV-specific and EV-unspecific characteristics might mediate the impact of these interventions on students' learning?

METHOD

Literature search and selection of studies

On April 26, 2022, we searched the high-ranked data bases Web of Science Core Collection, Scopus, Eric, PsycInfo, and Taylor & Francis Online Journals for the search terms diagram*, draw*, visual*, image*, sketch*, representation*, or graph* in the title, and math* in the whole text. In addition, we searched for peer-reviewed articles published between 2018 and 2022 in the English language for reasons of topicality and accessibility to the international community.

Our search identified 3,128 potentially relevant articles. To be included in the review, articles had to meet the following inclusion criteria: (a) the study focused on school mathematics learning or teaching with or through EV, (b) the study investigated an EV learning environment, (c) the study used a (quasi) experimental pre-posttest design or a posttest-only design to analyze the impact of an EV intervention on school students' mathematics learning. By screening Titles, Abstracts, and Keywords for inclusion criteria, we excluded 267 duplicates and 2,521 articles. Screening of the remaining full texts resulted in the exclusion of another 239 articles. We identified 41 articles that examined an EV intervention, out of which 12 articles met all the inclusion criteria. In the 12 articles, the authors investigated 17 EV learning environments by contrasting them against conventional learning or another EV learning environment.

Data extraction and analysis

We applied a qualitative content analysis to the full texts to systematically extract data on (1) reported efficacy, (2a) EV-unspecific characteristics of the learning environments, and (2b) EV-specific characteristics of the learning environments.

To describe the reported efficacy, we deductively coded *cognitive dependent measures* with the characteristics *understanding* and *problem solving* (including mathematical modelling) and inductively added *perception*, *interpretation*, and *mental rotation*. In addition, we coded whether taking part in the EV intervention had a *positive*, *zero*, or *negative* effect on learning compared with the comparison condition.

To extract EV-unspecific characteristics of the interventions, we first deductively applied the categories *school level*, *content domain*, and *intervention duration.* As EV-specific characteristics, we coded the *visualization component* (with the characteristics *process* and *product*) and *tool use* (with the characteristics *paper-pencil*, *technology*, *gestures*, and *hands-on objects*), on the basis of a previously developed coding scheme

(Schoenherr & Schukajlow, under review). To identify further characteristics, we applied a deductive-inductive procedure by first recording and then clustering characteristics mentioned by the authors into categories and assigning them to the EV-specific or EV-unspecific type. In this way, we added the EV-unspecific characteristics *metacognitive reflection on learning*, *scaffolding of learning*, and *student control* and the EV-specific characteristics *multiple EVs*, *visual interaction*, *metacognitive reflection on EV*, *scaffolding of EV*, and *EV transfer*.

As an indicator of coding reliability, two coders independently coded 25% of the included EV interventions on reported efficacy and EV-unspecific and EV-specific characteristics with a substantial percentage of agreement between 67% and 100%.

RESULTS

Efficacy of EV interventions

Comparing EV interventions with conventional learning, seven out of 12 EV interventions had a positive effect on student understanding and problem solving (e.g., Bernard & Senjayawati, 2019; Chen, 2019; Ke, 2019). For example, sixth and seventh graders who played an architecture simulation game including schematic EV outperformed students who were exposed to conventional learning in a problem-solving test on ratio, proportion, and area. Five EV interventions did not increase student learning compared with conventional learning (Ott, 2020; Rellensmann et al., 2021; Schoevers et al., 2020). As one example, providing students with an EV intervention on characteristics of accurate drawings did not result in increased modelling performance in geometry (Rellensmann et al., 2021). In studies comparing different EV interventions, findings were mixed with two studies reporting a positive effect (Aldalalah et al., 2019; Liang & She, 2021), one study reporting a positive effect for high-achieving students (Lee et al., 2018), and two studies reporting a null effect (Rellensmann et al., 2021; Soni & Okamoto, 2020). For example, Soni and Okamoto (2020) found that using number lines in a digital math game or in a paper-pencil workbook were equally effective at helping students learn fractions. Regarding visual perception, one study reported a positive effect after geometry training (Schoevers et al., 2020). No effects were found for non-geometry graphic interpretation (Lowrie et al., 2019) and mental rotation tasks (Ke, 2019; Ke & Clark, 2020).

Characteristics of effective EV interventions

On the *school level*, the majority of EV interventions addressed secondary school students (n = 10). The *duration* of EV interventions differed widely from four sessions of 15 min (Soni & Okamoto, 2020) to nine sessions of 60 to 90 min (Schoevers et al., 2020). The predominant *content domain* targeted in the EV interventions was geometry (n = 8), but other topics—for example, algebra (n = 3), probability (n = 1), and fractions (n = 1)—were also addressed. As we found positive and null effects across these characteristics, we cannot develop a conclusive hypothesis about the significance of the EV-unspecific characteristics educational level, intervention duration, and content domain for the efficacy of the EV interventions in this review.

A coding of the *visualization component* indicated that most studies (*n* = 9) included *visualization processes*. As one example, Lowrie et al. (2019) encouraged students in their EV intervention to mentally transform and manipulate 2D and 3D objects. Two studies that exclusively addressed *visualization products* did not find a positive effect on student learning (Ott, 2020; Rellensmann et al., 2021). For example, Ott (2020) addressed the product component by encouraging third-grade students to reflect on ready-made drawings in class. This led us to derive the hypothesis that addressing the visualization process component in EV interventions (i.e., all physical and mental activities and processes related to selecting, constructing, using, and interpreting EVs) is an important EV-specific characteristic for their efficacy.

In this review, seven studies used (amongst others) *technology* as a *tool* to construct or use EVs, two studies used hands-on objects, and four studies used paper-pencil only. Examples of technology used are Augmented Reality learning on mobile devices (e.g., Chen, 2019), an architecture simulation game (e.g., Ke & M. Clark, 2020), and dynamic geometry software (e.g., Lowrie et al., 2019). All studies using technology reported a positive effect on student learning, indicating that technology use might be an important EV-specific characteristic of effective EV interventions.

In addition, we extracted three EV-unspecific and six EV-specific characteristics that were considered potentially effective: The EV-unspecific characteristics consisted of *scaffolding of learning* (*n* = 3; e.g., Soni & Okamoto, 2020), *student control of learning* (i.e., individual learning pace and difficulty levels; *n* = 3; e.g., Chen, 2019), and *metacognitive reflection on learning*, including reflection on mathematical content, procedures, knowledge, and skills (e.g., Schoevers et al., 2020).

One frequently mentioned EV-specific characteristic (*n* = 7) was that the learning environment forced students to transfer information between *multiple EVs* (e.g., Bernard & Senjayawati, 2019; Liang & She, 2021), including concrete (Lowrie et al., 2019) and symbolic representations (Liang & She, 2021). Another EV-specific characteristic was *visual interaction* (*n* = 5), that is, the learning environment enabled students to visually observe, elaborate, explore, manipulate, and transform EVs (e.g., Ke, 2019). In addition, the authors proposed *metacognitive reflection on EV* (*n* = 2; Ott, 2020; Rellensmann et al., 2021), *scaffolding of EV* (*n* = 3; Ke & Clark, 2020), and *transfer of EV* across tasks (*n* = 3; e.g., Rellensmann et al., 2021) as promising factors that might increase the efficacy of EV interventions.

DISCUSSION

Of the 130 studies on EV in mathematics education research, a small proportion of studies (*n* = 12) used experimental designs to investigate EV interventions in schools. Our review of these studies showed a mixed—but mostly positive—impact on student learning in different mathematical topics, underlining the theoretically assumed benefit of EV as a medium for mathematics thinking and learning (e.g., Arcavi, 2003). The small number of experimental intervention studies indicates that more experimental studies are needed to obtain evidence for the effects of EV interventions on student

learning. Still, evidence of efficacy depends strongly on the choice of control condition and outcome measures. Findings indicate that EV interventions might be particularly effective in comparison with conventional learning (i.e., without EV) with respect to near transfer tasks that measure students' understanding of the learning topic (e.g., Soni & Okamoto, 2020).

Mixed findings on the efficacy of EV interventions have highlighted the need to gather information on the characteristics of effective EV interventions. An important novel contribution of this review is that we identified three EV-unspecific and six EV-specific intervention characteristics that might influence the efficacy of EV interventions.

Theoretically, our findings on the EV-specific intervention characteristics contribute to the framework of EV in mathematics education, as they point to key characteristics of EV in learning environments. As such, the EV process component (i.e., learning how to construct, generate, use, and interpret EV; Arcavi, 2003) seems to be important for learning with or through EV.

Empirically, the characteristics we identified confirm and add to previously identified characteristics from EV research and other research areas. For example, our analyses supported the previously identified characteristics *metacognitive reflection on learning* (Dignath & Büttner, 2008), *student control of learning and multiple EVs* (Duijzer et al., 2019), *metacognitive reflection on EV* and *scaffolding of EV* (Fiorella & Zhang, 2017), and *transfer* of EV to different tasks to help students generalize characteristics of specific EV tasks (Presmeg, 1986). An important new contribution is that our analyses also uncovered the EV-specific characteristics *visualization process*, *technology use*, and *visual interaction*. This means, for example, that we hypothesize that providing students with opportunities to visually explore, manipulate, and transform EVs will increase student learning in learning environments that focus on EV. Further research is needed to determine the differential impact of the characteristics in order to contribute to a better understanding of how they influence the efficacy of EV interventions. As one example, technology use appeared to be positively related to the intervention's efficacy, a finding that might be explained by the use of individual learning sessions that provided students with scaffolding and opportunities for visual interaction.

Practically, as most of the characteristics have been supported by prior research, the EV-specific and EV-unspecific characteristics we identified can help practitioners design EV learning environments.

Limitations

In this review, we applied an extensive automatic search strategy to identify a wide range of studies that investigated EV interventions. Still, we might have missed some relevant studies that may have been framed differently. Our analysis of the learning environments was based on the information provided in the papers. To increase the objectivity of our coding, we relied on the terms used by the authors whenever possible

(e.g., problem solving). However, different interpretations by the authors might bias this review's findings. Also, we analyzed a wide variety of learning environments to gain insights into EV intervention research and to determine the extent to which different interventions support student learning. For these reasons, it is difficult to generalize our results, and more research is needed on the benefits and boundaries of learning with or through EV. In this review, we focused on intervention characteristics. In addition, recent research has pointed to learner characteristics that influence the impact of EV interventions (e.g., mathematical abilities; Lee et al., 2018). More research is needed on learner characteristics and their interplay with intervention characteristics in promoting students' learning in EV learning environments.

REFERENCES

Papers included in this review have been highlighted (*).

*Aldalalah, O., Ababneh, Z. W., Bawaneh, A. K., & Alzubi, W. M. (2019). Effect of augmented reality and simulation on the achievement of mathematics and visual thinking among students. *International Journal of Emerging Technologies in Learning*, *14*(18), 164–185. https://doi.org/10.3991/ijet.v14i18.10748

Arcavi, A. (2003). The role of visual representations in the learning of mathematics. *Educational Studies in Mathematics*, *52*(3), 215–241. https://doi.org/10.1023/A:1024312321077

*Bernard, M., & Senjayawati, E. (2019). Developing the students' ability in understanding mathematics and self-confidence with VBA for Excel. *Journal of Research and Advances in Mathematics Education*, *4*(1), 45–56. https://journals.ums.ac.id/index.php/jramathedu/article/view/6349/4440

*Chen, Y. (2019). Effect of mobile augmented reality on learning performance, motivation, and math anxiety in a math course. *Journal of Educational Computing Research*, *57*(7), 1695–1722. https://doi.org/10.1177/0735633119854036

Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning*, *3*(3), 231–264. https://doi.org/10.1007/s11409-008-9029-x

Duijzer, C., van den Heuvel-Panhuizen, M., Veldhuis, M., Doorman, M., & Leseman, P. (2019). Embodied learning environments for graphing motion: A systematic literature review. *Educational Psychology Review*, *31*(3), 597–629. https://doi.org/10.1007/s10648-019-09471-7

Fiorella, L., & Zhang, Q. (2018): Drawing boundary conditions for learning by drawing. *Educational Psychology Review*, 30(3), 1115–1137. https://doi.org/10.1007/s10648-018-9444-8

*Ke, F. (2019). Mathematical problem solving and learning in an architecture-themed epistemic game. *Educational Technology Research and Development*, *67*(5), 1085–1104. https://doi.org/10.1007/s11423-018-09643-2

*Ke, F., & M. Clark, K. (2020). Game-based multimodal representations and mathematical problem solving. *International Journal of Science and Mathematics Education*, *18*(1), 103–122. https://doi.org/10.1007/s10763-018-9938-3

*Lee, C.-Y., Lei, K. H., Chen, M.-J., Tso, T.-Y., & Chen, I.-P. (2018). Enhancing understanding through the use of structured representations. *Eurasia Journal of Mathematics, Science and Technology Education*, *14*(5), 1875–1886. https://doi.org/10.29333/ejmste/85424

*Liang, C.-P., & She, H.-C. (2021). Investigate the effectiveness of single and multiple representational scaffolds on mathematics problem solving: Evidence from eye movements. *Interactive Learning Environments.* https://doi.org/10.1080/10494820.2021.1943692

*Lowrie, T., Logan, T., & Hegarty, M. (2019). The influence of spatial visualization training on students' spatial reasoning and mathematics performance. *Journal of Cognition and Development*, *20*(5), 729–751. https://doi.org/10.1080/15248372.2019.1653298

OECD. (2019). *Pisa 2018 Assessment and Analytical Framework*. *PISA*. OECD Publishing. https://doi.org/10.1787/b25efab8-en

*Ott, B. (2020). Learner-generated graphic representations for word problems: An intervention and evaluation study in grade 3. *Educational Studies in Mathematics*, *105*(1), 91–113. https://doi.org/10.1007/s10649-020-09978-9

Presmeg, N. (1986): Visualisation in high school mathematics. *For the Learning of Mathematics*, 6 (3), 42–46. http://www.jstor.org/stable/40247826.

*Rellensmann, J., Schukajlow, S., Blomberg, J., & Leopold, C. (2021). Does strategic knowledge matter? Effects of strategic knowledge about drawing on students' modeling competencies in the domain of geometry. *Mathematical Thinking and Learning.* Advance online publication. https://doi.org/10.1080/10986065.2021.2012741

Schnotz, W. (2005). An integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 49–69). Cambridge University Press.

Schoenherr, J., & Schukajlow, S. (under review). Characterizing external visualization in mathematics education research: A scoping review. *ZDM – Mathematics Education*.

*Schoevers, E. M., Leseman, P. P., & Kroesbergen, E. H. (2020). Enriching mathematics education with visual arts: Effects on elementary school students' ability in geometry and visual arts. *International Journal of Science and Mathematics Education*, *18*(8), 1613–1634. https://doi.org/10.1007/s10763-019-10018-z

*Soni, M., & Okamoto, Y. (2020). Improving children's fraction understanding through the use of number lines. *Mathematical Thinking and Learning*, *22*(3), 233–243. https://doi.org/10.1080/10986065.202

TEACHERS' DIAGNOSTIC ACTIVITIES DURING TASK-BASED ASSESSMENTS IN A DIGITAL SIMULATION

Christian Schons[1], Andreas Obersteiner[1], Frank Fischer[2] and Kristina M. Reiss[1]

[1]Technical University of Munich, Germany

[2]Ludwig-Maximilians-University of Munich, Germany

Teachers need to be able to assess students' learning outcomes accurately. While teachers' assessment competencies have long been conceptualised solely as accuracy of assessments, recent research also focusses on situation-specific processes. However, conceptualising and assessing such processes reliably is still a challenge. We analysed 76 pre-service teachers' diagnostic activities during task-based assessments in a digital simulation. Categories of diagnostic activities were derived from a model of scientific reasoning and argumentation. Participants mostly evaluated the available evidence and rarely stated hypotheses or drew conclusions. There were pronounced individual differences in the frequencies of these activities. The results provide a basis for individualising support of pre-service teachers' diagnosing.

THEORETICAL BACKGROUND

Assessing individual students' learning outcomes is a challenging task for teachers (Südkamp et al., 2012). While many studies focused on the quality of teachers' assessments (most prominently, their accuracy), there is still a lack of understanding of situation-specific assessment processes in complex assessment situations – which is necessary to foster assessment competencies effectively (Heitzmann et al., 2019; Loibl et al., 2020). In mathematics lessons, teachers often assess their students' mathematical competencies by interpreting students' solutions to mathematical tasks. While several studies focused on teachers' task selections in learning situations, only few empirical studies investigated teachers' competencies in task-based assessment situations. In task-based assessments, teachers should be able to evaluate students' task solutions appropriately (e.g., to notice students' errors) and to draw the right conclusions (e.g., interpret any systematic errors in terms of the students' underlying misconception). Currently, there is still a need for conceptualising and measuring reliably situation-specific processes during task-based assessments, and little is known about (pre-service) teachers' individual differences in these processes.

In this study, we analysed pre-service mathematics teachers' assessment processes in a task-based assessment situation, using a digital simulation. We aimed at 1) describing pre-service teachers' diagnostic activities during the interpretation of students' task solutions based on a model of scientific reasoning and argumentation and 2) identifying different groups of pre-service teachers based on these activities.

2023. 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. 171-178). PME 46.

Teachers' Assessment Competencies

Teachers need assessment competencies for accurately assessing students' outcomes (Leuders et al., 2022). Teachers' assessment competencies include individual dispositions (e.g., professional knowledge) and situation-specific processes (i.e., the actual behaviour in assessment situations) that influence performance in assessment situations (Blömeke et al., 2015; Heitzmann et al., 2019). In contrast to professional knowledge, which can be measured by paper-pencil tests, measuring situation-specific processes requires practical challenges that allow for observation of behavioural processes. Several studies have focused on such processes during assessments of task features, while only few studies investigated more complex assessment situations.

Research on Situation-Specific Assessment Processes

Some researchers investigated teachers' ability to perceive and interpret task difficulties (e.g., Rieu et al., 2022) or instructional features of tasks (e.g., Schreiter et al., 2021). These studies manipulated participants' professional knowledge in experimental studies, and they conceptualised assessment processes as cognitive information processing (Loibl et al., 2020). Results indicated that direct instruction of specific pedagogical content knowledge had a positive impact on pre-service teachers' perceiving and interpreting of task features (Rieu et al., 2022) in various content domains (e.g., fractions and functions). Other studies described teachers' assessment processes in terms of eye-movements (e.g., Schreiter et al., 2021). Fixation frequencies and fixation durations were used as indicators of efficiency of information processing, and inducing specific pedagogical content knowledge caused higher efficiency.

In classroom situations, teachers not only need to analyse task features, but also to assess students' mathematical competencies. Studies that investigated teachers' assessments of students' mathematical competencies used video-vignettes of students in authentic situations (e.g., Sommerhoff et al., 2023). They analysed pre-service teachers' reasoning about students' mathematical competencies by coding participants' written notes that were recorded during a simulation. Results indicated that pre-service teachers mostly described superficial observations of students' solutions, whereas they rarely integrated pieces of evidence to draw conclusions for getting a larger picture of students' competencies. Other studies used qualitative interviews to explore teachers' interpreting of students' task solutions (Philipp & Leuders, 2014; Son, 2013). One finding was that (pre-service) teachers differed in their argumentation skills and their ability to create artefacts (e.g., mathematical tasks) in order to test specific hypotheses about student thinking.

For investigating more complex assessment situations, concepts from the theory of scientific reasoning and argumentation that occur in task-based assessment situations are promising to complement existing approaches to analyse situation-specific assessment processes (Fischer et al., 2014; Heitzmann et al., 2019).

Diagnostic Activities

A framework for scientific reasoning and argumentation by Fischer et al. (2014) describes a set of epistemic activities that people engage in while creating knowledge. Making assessments based on task solutions includes a process of creating knowledge about students' mathematical competencies and Heitzmann et al. (2019) specified the epistemic activities for assessment situations (*diagnostic activities*). The following activities are likely to occur during task-based assessments:

Evidence Evaluation. A teacher describes characteristics of one specific task solution.

Stating Hypotheses. A teacher states a hypothesis about a student's competencies or systematic errors based on several task solutions.

Drawing Conclusions. A teacher draws a conclusion about a student's competencies or systematic errors based on several task solutions.

There is yet little empirical evidence about how often pre-service teachers engage in these diagnostic activities during task-based assessments and which individual differences can be observed among pre-service teachers.

RESEARCH QUESTIONS

In this study, we aim at analysing pre-service teachers' diagnostic activities during task-based assessments of students' mathematical competencies in a digital simulation. We also aim at identifying individual differences in the occurrence of these activities among pre-service teachers. The specific research questions were:

1) Can we measure pre-service teachers' diagnostic activities from their written notes they make during task-based assessments in a digital simulation with sufficient reliability? 2) Which diagnostic activities do pre-service teachers engage in during task-based assessments? 3) Are there different groups of pre-service teachers based on their patterns of diagnostic activities?

To address the first research question, we evaluated a coding scheme for assigning participants' written notes in a digital simulation to the categories *Evidence Evaluation*, *Stating Hypotheses* and *Drawing Conclusions* (Heitzmann et al., 2019). For the second research question, we expected that pre-service teachers mostly evaluate evidence and rarely state hypotheses or draw conclusions (Sommerhoff et al., 2023). We investigated the third research question on an explorative basis, expecting individual differences in pre-service teachers' diagnostic activities.

METHODS

Sample

Participants were 76 pre-service primary school mathematics teachers (61 female) studying at a German university. Their mean age was 21.60 years (SD = 2.53) and their median semester of university study was three (IQR = 3–2).

The Digital Simulation

To examine participants' diagnostic activities, we used a digital simulation (Wildgans-Lang et al., 2022) that provided a task-based assessment situation. Based on mathematical tasks and its real solutions, the digital simulation provided different student cases differing in their competencies and systematic errors.

The Tasks. Tasks were developed within a framework of mathematical competencies of third graders and validated from a theoretical and empirical perspective. Their real solutions were taken from a pilot run of a national large-scale assessment among third-graders in Germany (VERA-3). The tasks varied in their difficulty levels and in their potential to reveal students' systematic errors.

Simulated Student Cases. The simulation provided six different student cases. Each student case consisted of a set of tasks and corresponding solutions, some of which showed typical errors. The student cases varied in their mathematical competencies and the tasks of each student case included one systematic error (e.g. in multi-digit subtractions, see Figure 1). To assess one student case in the digital simulation, participants first selected a blank task, whereupon the corresponding student's solution appeared. After interpreting a solution, participants could select another task to generate more information about the same student case. Participants could repeat this process multiple times and terminate anytime in order to come to a final conclusion. During the presentation of a task solution, a text box was shown in which participants took notes that were recorded in the log data (see Figure 1).

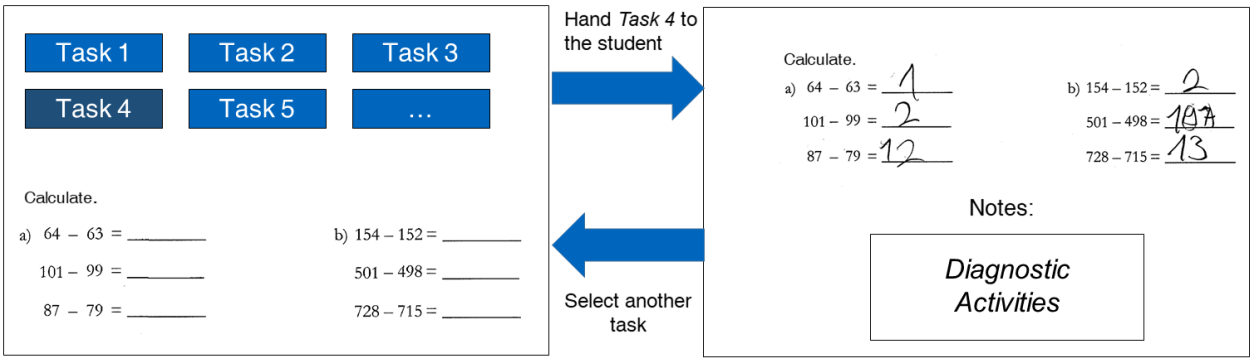


Figure 1: Overview of the assessment of a student case in the digital simulation with an exemplary task and the corresponding student solution.

Procedure

Participants took part in this study within a 90-minute seminar session which was part of their regular curriculum. Before the session, participants were informed about the aim of their assessments, that is, to assign student cases to a competence level and to identify their systematic errors. Participants were instructed to continue as long with the assessment of a student case until they were sure about their assessment. They were instructed to note their thoughts in the text box provided while they were assessing.

Coding of Diagnostic Activities

For analysing participants' written notes during their assessments, we derived three activities that likely occur during the interpretation of task solutions from Heitzmann et al.'s (2019) model of diagnostic activities. Table 1 provides an overview of the categories and typical examples.

Category	*Explanation*	*Typical Example*
Evidence Evaluation	Description of apparent features in one of a student's solutions. Statement solely refers to salient features of a solution.	„S1 is adding instead of multiplying“ „Dividing by 4 is correct“
Stating Hypotheses	Stating hypotheses about a student's potential systematic error. Statement contains an expression of uncertainty.	„S2 interchanged 340 and 430, which might indicate difficulties in shifting between spoken and written numbers.”
Drawing Conclusions	Drawing conclusions about student's systematic error. Statement contains no expression of uncertainty.	„She lacks understanding of the decimal system“ „The student has many errors because of a misconception of the place-value-system“
Default	Notes that cannot be assigned any category	„?“

Table 1: Summary of the coding scheme with typical examples for the categories *Describing*, *Stating Hypotheses* and *Drawing Conclusions*.

Data Analysis

For the identification of different groups of pre-service teachers' based on their diagnostic activities, participants were clustered on their activities in the digital simulation using the *k*-means algorithm. To this end, standardised mean scores were used. First, we determined the optimal number of clusters using *Euclidean distance* and the *Ward method*. We continued our analyses with the number of clusters suggested by the majority of the stopping rules to build the final cluster structure by assigning the participants in our sample to the determined number of clusters.

RESULTS

Overall, participants made 313 assessments in the digital simulation and took 3458 notes during their assessments. Two research assistants applied the coding scheme (see Table 1) to pre-service teachers' written notes and had substantial agreement (Cohen's

κ = .76). This means that we succeeded in reliably assessing participants' diagnostic activities based on their written notes.

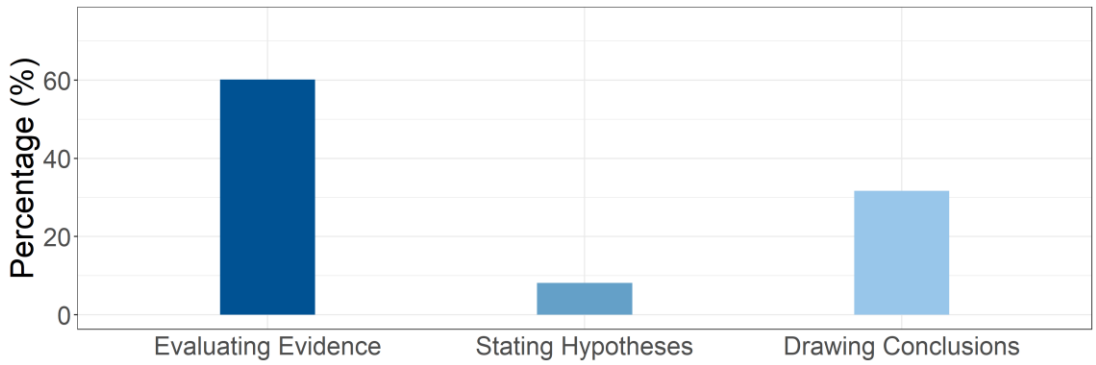


Figure 2: Relative frequencies of diagnostic activities during assessment processes in the digital simulation across all participants

Regarding the diagnostic activities participants performed during the interpretation of students' task solutions, we found that participants mostly evaluated evidence (60.1% of the coded activities) and drew conclusions (31.7%), whereas they rather rarely stated hypotheses (8.1%, see Figure 2).

The third question asked whether we could identify groups of pre-service teachers based on their activities. According to the majority rule, the number of clusters that represented the data best was three. We describe these three clusters in relation to the whole sample (see Fig. 2): The first cluster (*baseline*) consisted of 52 pre-service teachers who had comparably few activities of all three types while making assessments. The second cluster (*drawing conclusions*) consisted of 10 pre-service teachers who relatively often drew conclusions about students' competencies and systematic errors, but rarely evaluated evidence or stated hypotheses during the assessment process. The 14 pre-service teachers in the third cluster (*stating hypotheses*) often evaluated evidence or stated hypotheses, but rarely drew conclusions.

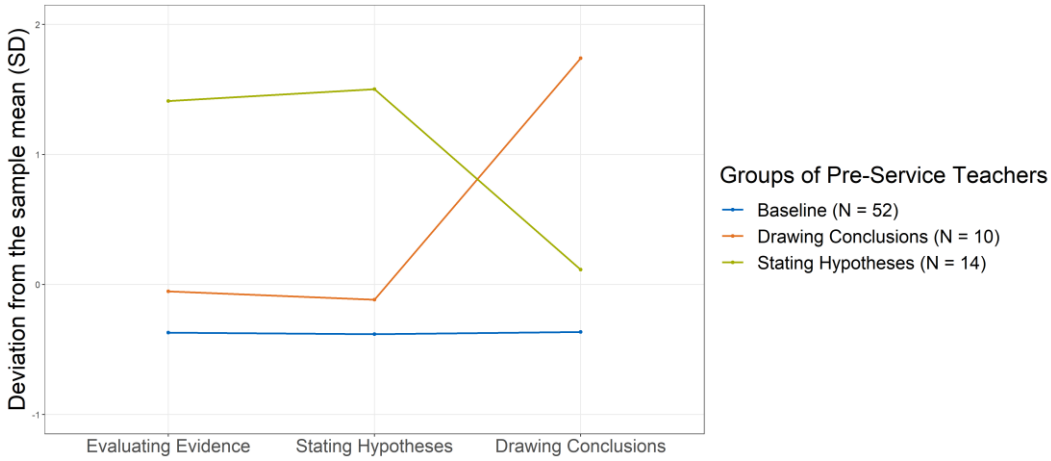


Figure 2: Cluster centers of the three identified clusters resulting from a cluster analysis of 76 pre-service teachers based on their diagnostic activities.

DISCUSSION

This study aimed at analysing pre-service teachers' situation-specific assessment processes in task-based assessments, conceptualised as diagnostic activities. To this end, we developed a coding scheme to measure diagnostic activities based on pre-service teachers' written notes in the simulation. We succeeded in reliably coding these notes. In line with the results reported by Sommerhoff et al. (2023), participants mostly engaged in describing features of students' solutions to mathematical tasks, and rarely integrated different pieces of evidence to get a larger picture of students' mathematical competencies. Going beyond previous studies, we identified pronounced individual differences among our sample. This might have implications for including instructional support measures to foster teachers' assessment competencies in the simulation. Specifically, the majority of participants (those in the "baseline" and the "drawing conclusions" cluster) may benefit from prompts that encourage them to state hypotheses about students' competencies and misconceptions, which could make their diagnosing more systematic. Participants in the "stating hypotheses" cluster, on the other hand, may benefit from prompts that encourage them to draw conclusions, based on their observed evidence and their hypotheses. Preliminary data from a follow-up study suggest that such prompts may lead to more efficient diagnostic processes and, eventually, more accurate diagnoses.

A limitation of the present study is the measurement of diagnostic activities based on their written notes, because participants may not have noted all their thoughts. Moreover, expanding the analysis to include qualitative categories might provide a quality measure of teachers' assessment competencies beyond accuracy.

ACKNOWLEDGEMENTS

This research is part of the COSIMA research group, funded by German Research Foundation (FOR 2385, Project: 5150171 DG RE 1247/12-2)

REFERENCES

Blömeke, S., Gustafsson, J.-E., & Shavelson, R. J. (2015). Beyond Dichotomies. *Zeitschrift Für Psychologie*, *223*(1), 3–13.

Fischer, F., Kollar, I., Ufer, S., Sodian, B., Hussmann, H., Pekrun, R., Neuhaus, B., Dorner, B., Pankofer, S., Fischer, M., Strijbos, J.-W., Heene, M., & Eberle, J. (2014). Scientific Reasoning and Argumentation: Advancing an Interdisciplinary Research Agenda in Education. *Frontline Learning Research*, *2*(3), 28–45.

Heitzmann, N., Seidel, T., Opitz, A., Hetmanek, A., Wecker, C., Fischer, M., Ufer, S., Schmidmaier, R., Neuhaus, B., Siebeck, M., Stürmer, K., Obersteiner, A., Reiss, K., Girwidz, R., & Fischer, F. (2019). Facilitating Diagnostic Competences in Simulations: A Conceptual Framework and a Research Agenda for Medical and Teacher Education. *Frontline Learning Research*, *7*(4), 1–24.

Loibl, K., Leuders, T., & Dörfler, T. (2020). A Framework for Explaining Teachers' Diagnostic Judgements by Cognitive Modeling (DiaCoM). *Teaching and Teacher Education*, *91*, 103059.

Philipp, K., & Leuders, T. (2014). Diagnostic competences of mathematics teachers – processes and resources. In P. Liljedahl, S. Oesterle, C. Nicol, & D. Allan (Eds.), *Proceedings of the 38th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 425–432). PME.

Rieu, A., Leuders, T., & Loibl, K. (2022). Teachers' diagnostic judgments on tasks as information processing – The role of pedagogical content knowledge for task diagnosis. *Teaching and Teacher Education*, *111*, 103621.

Schreiter, S., Vogel, M., Rehm, M., Dörfler, T. (2021). Through the eyes of prospective teachers: Judging task difficulties in the domain of fractions. 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. 191-198). PME.

Sommerhoff, D., Codreanu, E., Nickl, M., Ufer, S., & Seidel, T. (2023). Pre-service teachers' learning of diagnostic skills in a video-based simulation: Effects of conceptual vs. interconnecting prompts on judgment accuracy and the diagnostic process. *Learning and Instruction*, *83*, 101689.

Son, J.-W. (2013). How preservice teachers interpret and respond to student errors: Ratio and proportion in similar rectangles. *Educational Studies in Mathematics*, *84*(1), 49–70.

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.

Wildgans-Lang, A., Scheuerer, S., Obersteiner, A., Fischer, F., & Reiss, K. (2022). Learning to Diagnose Primary Students' Mathematical Competence Levels and Misconceptions in Document-Based Simulations. In: Fischer, F., Opitz, A. (Eds.) Learning to Diagnose with Simulations. Springer.

STATISTICAL THINKING AND VIEWING PATTERNS WHEN COMPARING DATA DISTRIBUTIONS: AN EYE-TRACKING STUDY WITH 6TH AND 8TH GRADERS

Saskia Schreiter and Markus Vogel

Heidelberg University of Education, German

Many students tend to perceive a data distribution as a collection of individual values rather than as a conceptual entity (local vs. global view of data). These difficulties seem to persist even after instruction in statistics. This study uses a methodological triangulation of eye-tracking and stimulated recall interviews to examine and contrast 6th and 8th grade students' (N = 49) viewing patterns and statistical thinking when comparing data distributions. Results showed no significant differences between 6th and 8th graders. Regardless of students' grade level, the empirical data confirmed our theoretically derived hypotheses for differences in certain eye-tracking measures (fixation count, saccade amplitude, saccade direction) between students with a local and global view of data.

THEORETICAL BACKGROUND

Perspectives on data distributions: local vs. global view

Distribution comparisons provide motivating learning opportunities in schools to initiate statistical thinking already before formal procedures of inferential statistics are known (Konold & Higgins, 2003). Many studies use the context of distributional comparison to examine students' conceptions of data distributions and how they relate to their data-based decisions regarding distributional comparison (e. g., Ben-Zvi & Arcavi 2001; Frischemeier, 2019; Gal et al., 1989; Watson & Moritz, 1999). The results of these studies have shown that students often struggle to understand a data distribution as a conceptual entity. These difficulties seem to persist statistics instruction and are reflected in students' tendency to focus on local details of the distributions without paying attention to the differences between the two distributions as a whole (*local vs. global view of data*; Bakker & Gravemeijer, 2004; Ben-Zvi, 2004). Table 1 provides an overview of the global and local characteristics of distributions that can be considered when comparing data distributions. This framework is based on Bakker and Gravemeijer (2004) and has been adapted for the objectives of this study with an explicit focus on visually determinable features of distributions. The structure of the framework can be read in two directions: It is typical for the upward perspective (local view of the data) that students see the distribution as a collection of individual data points from which they can determine, for example, the mean, range, or quartiles. However, this does not necessarily mean that students view these characteristics as measures of mean and range or as representatives of a group (Konold & Pollatsek, 2004). Therefore, it is important that students also develop the downward perspective (global view of the data) that is considered essential for statistical data analysis (Ben-

2023. 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. 179-186). PME 46.

Zvi, 2004). While many studies have focused on students' statistical reasoning when comparing data distributions, little is known about the underlying perceptual and attentional processes that guide students in choosing or dismissing features.

Global view	**distribution** (conceptual entity)		*Local view*
	global features		
	center *visually estimated mean, modal value…*	spread/density *observed spread-out-ness, majority, range…*	shape *observed skewness, peak, tails…*
	local features		
	Single/multiple isolated data points *extreme values, outlier, absolute frequencies of dots in certain intervals, value of certain intervals of dots, disjoint edge values…*		

Table 1: Perspectives on data distributions: local vs. global view
(Schreiter & Vogel, in press)

Eye-Tracking measures as indicator for a local vs. global view of data

Numerous studies in mathematics education research have shown that eye-tracking has a high potential to provide new insights into students' mathematical thinking and learning (for an overview, see Strohmaier et al., 2020). Recent studies in statistics education illustrated that eye-tracking is an effective method to study students' strategies and difficulties when interpreting and comparing statistical graphs such as histograms (e.g., Boels et al., 2019)

In a first step of our ongoing research project, hypotheses for differences in certain eye-tracking measures (fixation count, saccade amplitude, saccade direction) between students with a local and global view of data were theoretically derived and empirically investigated with a sample of 25 6th grade students (Schreiter & Vogel, in press). The results confirmed our hypotheses by showing that students with a global compared to a local view of data had on average significantly fewer fixations, longer saccade amplitudes and a higher relative number of horizontal saccades. Figure 1 illustrates the meaning of these eye-tracking measures by two exemplary gaze plots of students: the gaze plot on the left side shows a high number of fixations, and many saccades in vertical direction with short amplitudes. These are indicators for a local view of the data distribution, where local features (i.e., individual data points) are perceived and processed. In comparison, the gaze plot on the right side shows a smaller number of fixations, and more saccades in the horizontal direction with longer amplitudes within the data distributions. These are indicators of a global view of the data distributions, where global features that relate to the distribution as a whole are perceived and processed.

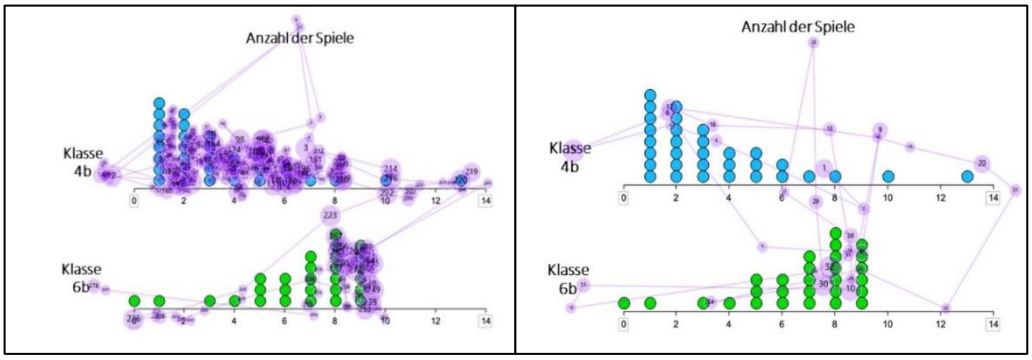


Figure 1: Exemplary gaze plots of two students comparing data distributions

In this paper, we pursue two research aims: First, we aim to replicate the findings of our study with 6th graders with empirical data from a sample of 8th graders. Our first research question is: *Is a global view (compared to a local view) in the 8th grade sample also associated with on average significantly fewer fixations, longer saccade amplitudes and a higher relative number of horizontal saccades? (RQ1)* Second, we aim to compare the 6th and 8th graders regarding potential differences in viewing behavior and statistical thinking. Our second, two-fold research question is: *Do 6th and 8th graders differ in terms of viewing behavior (RQ2a) and statistical thinking (with regard to the perception and processing of local vs. global distributional features) (RQ2b) when comparing data distributions?*

METHODS

Sample. The results reported here are based on data from n = 25 6th grade students (56% female) and n = 24 8th grade students (58% female). On average, 6th grade students were 11.6 years old (SD = 0.57) and 8th grade students were 13.6 years old (SD = 0.65). The students were recruited from three German secondary schools of type Gymnasium and Realschule. According to their curriculum, students were formally introduced to determine specific local distributional features (e.g., maximum, minimum) and global features of center (e.g., arithmetic mean) in grades 5/6, and to determine measures of spread (e.g., range, quartile) in grades 7/8.

Material. Four items on distribution comparisons were created. All items included authentic comparison situations and an explicit request to draw a conclusion from the data presented in the task. Between items, certain features in which the distributions differ (such as center, spread or shape) were varied. In addition, sample sizes were varied (one item with very different sample sizes, two items with slightly different sample sizes, and one item with equal sample sizes). This systematic variation was chosen to test whether students switch flexibly between local and global strategies depending on certain features of the distributional comparison. For example, while comparing absolute frequencies of dots in certain intervals is a valid local strategy to compare samples of equal sizes, it is incorrect to do so when sample sizes are unequal. To assess the comprehensibility of the items and study procedure, a pilot study was conducted with N = 20 6th grade students.

Procedure. The study was conducted at the children's schools. An example item was used to explain the procedure of the study. The four items were presented individually and in randomized order on a 24-inch computer screen (Fujitsu B24T-7 LED, 1920 × 1080 pixels). Eye-tracking data was collected using a monitor-based eye-tracker (Tobii Pro Fusion) that captured binocular eye movements at a sampling rate of up to 120 Hz. For adjusting the eye-tracker, a 9-point calibration was performed before each task. An eye-tracking stimulated recall interview was conducted directly following each task to avoid loss of memory. Eye-tracking stimulated recall interviews have shown to be an effective method to examine students' cognitive processes by asking them to retrospectively describe their own thoughts and actions as precisely as possible based on a video of their own gaze movements (Schindler & Lilienthal, 2019). For the recording of the interviews, the software OBS was used, which recorded screen contents including sound, so that the videos of the eye movements with the corresponding comments of the subjects were available for the later analysis.

Data analysis. Tobii Pro Lab software was used to analyze the eye-tracking data. In each task, two Areas of Interest (AOIs) were defined that covered the two distributions. Fixations and saccades within the AOIs were detected using the Tobii I-VT Fixation Filter. This filter classified eye movements either as part of a fixation if the velocity is below the threshold of 30°/s, or as part of a saccade if the velocity is equal to or higher than this threshold. To determine the saccadic measures (saccade amplitude, saccade direction), raw gaze data was analyzed. Only saccades that have the immediately preceding and consecutive fixations within the same AOI were considered for the analysis. Saccade direction was defined as the absolute angle of a saccade (in degrees) measured to the horizontal and calculated based on the coordinates of the immediately preceding and consecutive fixations using basic trigonometry (cf. Holmqvist & Andersson, 2017). Saccades were classified as horizontal if this angle was between 0 – |44|° and classified as vertical if it was between |45|°– |90|°. In total, four recordings had to be excluded due to data loss or insufficient accuracy.

The data of the stimulated recall interviews was coded both deductively and inductively using qualitative content analysis. As the focus of the research presented here was on students' statistical thinking while comparing data distributions, the coding procedure only referred to the process until the distribution comparison decision was made. Students' utterances were analyzed in terms of perceived and processed distributional features and were assigned to two main categories (local feature/global feature) and several sub-categories (based on the framework presented in table 1). In total 196 videos were coded. Seven recordings were excluded due to technical problems with audio or screen recording. All transcripts were coded independently by two raters with high interrater reliability (Cohen's kappa = .84).

RESULTS

Statistical thinking when comparing data distribution

To gain information on students' statistical thinking with regard to the perception and processing of global and local distributional features, we analyzed the stimulated recall interviews in a qualitative manner. Figure 2 gives an overview of the results regarding the relative number of items in which at least one global feature was perceived and processed. The majority of the 6th grade students (68%) and about half (52%) of the 8th grade students did either not consider global features in any of the items or in all four items (figure 2). The students who did not consider any global characteristics across all items are to be classified as problematic in this context. These students remained with local strategies (e.g., comparing absolute frequencies of dots in certain intervals or comparing the value of certain intervals of dots), even if sample sizes are unequal, which is an incorrect strategy in these cases. Overall, results showed that there was no statistically significant difference between the 6th graders and 8th graders in terms of the relative number of items in which at least one global feature was perceived and processed, $t(46) = 9.41$, $p = .352$.

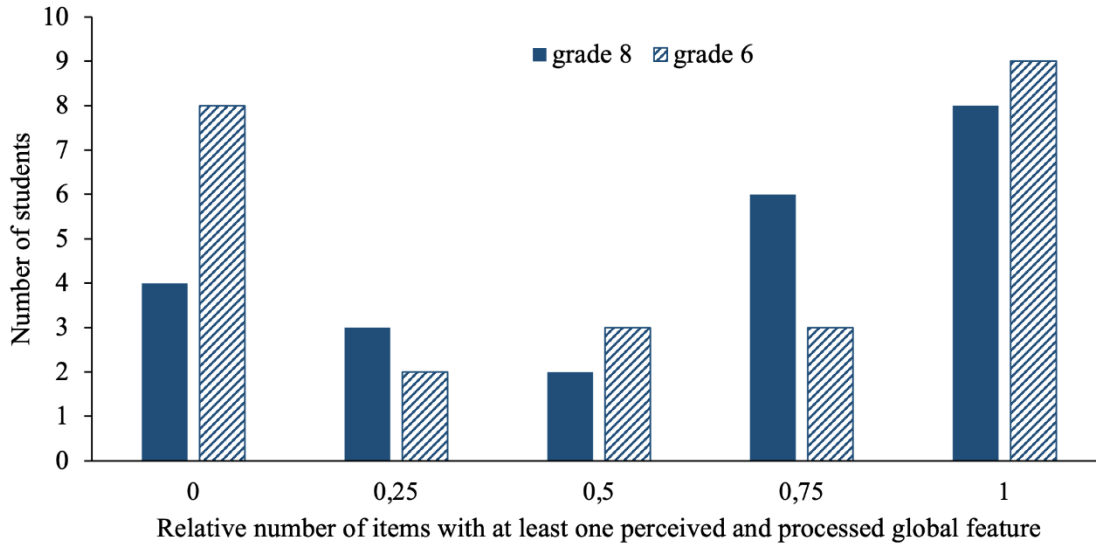


Figure 2: Students' statistical thinking in terms of the relative number of items in which at least one global feature was perceived and processed

Regarding global features, students' utterances were assigned to the three categories center, spread/density, and shape. For both 6th and 8th graders, the spread/density category was often assigned when students compared areas with particularly many dots (so-called modal clumps) or divided the distributions into three groups (of low, middle, and high values). Sometimes, students also referred to range or compared how "spread out" or how "close together" the data points of the distribution are. Regarding center, it was observed among both grade levels that students mostly identified and compared the modal values of the two distributions. Visual estimation of the arithmetic mean rarely took place. Regarding shape, students' utterances were very different, comparing

the shape of the distributions for example to geometric shapes like triangles or pyramids or to objects from their everyday life such as stairs that go up and down.

Viewing patterns when comparing distributions

Based on the analysis of the stimulated recall interviews, the sample was split in those students who perceived and processed at least one global feature in half or more of the items (from now on referred to as students with a global view) and those students below that threshold (from now on referred to as students with a local view). Table 2 provides an overview of the collected eye-tracking measures separated by students with a local and global view and by 6th and 8th graders. To address potential group differences in the collected eye-tracking measures, a MANOVA with the two between subject factors *view of data* (local/global) and *grade level* (6/8) was calculated.

Results showed a statistically significant difference between students with a local and global view on the combined dependent variables, $F(3, 42) = 7.827$, $p < .001$, partial $\eta^2 = .359$, Wilk's $\Lambda = .641$. This effect is independent of students' grade level (no significant interaction between *view of data* and *grade level*, $p = .872$). Regardless of grade level, students with a global compared to a local view of data showed on average significantly fewer fixations, longer saccade amplitudes, and a higher relative number of horizontal saccades. These effects were highly significant (all $p < .01$) with high effect sizes (all $\eta^2 > .20$).

No significant effect was found between 6th and 8th graders on the combined dependent variables, $F(3, 42) = 1.157$, $p = .338$, partial $\eta^2 = .076$, Wilk's $\Lambda = .924$.

ET measures	Local		Global	
	Grade 6 *M* (*SD*)	Grade 8 *M* (*SD*)	Grade 6 *M* (*SD*)	Grade 8 *M* (*SD*)
Number of fixations	232.63 (144.66)	173.86 (83.39)	95.58 (67.64)	82.53 (56.51)
Saccade amplitude (pix.)	78.12 (17.43)	96.35 (46.12)	124.68 (30.85)	132.84 (29.40)
Rel. number of horiz. saccades	0.51 (0.09)	0.54 (0.15)	0.63 (0.08)	0.62 (0.09)

Table 2: Descriptive statistics for eye-tracking measures of students with a local (grade 6: $n = 10$, grade 8: $n = 7$) and global view (grade 6: $n = 15$, grade 8: $n = 16$)

DISCUSSION

This study investigated 6th and 8th grade students' visual attention and statistical thinking while comparing data distributions.

With regard to our first research question, we analyzed if the findings of the 6th grade sample can be replicated with empirical data of an 8th grade sample. Results showed that the differences in certain eye-tracking measures (fixation count, saccade

amplitude, saccade direction) between students with a local and global view of data existed regardless of their grade level. In line with our theoretically derived hypotheses (Schreiter & Vogel, in press), both 6th and 8th grade students had on average significantly fewer fixations, longer saccade amplitudes, and a higher relative number of horizontal saccades with a global compared to a local view of data. These results suggest that eye-tracking data can help to identify students' conceptions and difficulties related to a local vs. global view of data. Knowledge about what features students focus their visual attention on and what is going on in students' minds as they visually focus on these features may provide further insight into how tasks and instruction should be designed to guide students from a local to a global view on data, which is considered an important goal of statistics education (Ben-Zvi & Arcavi, 2001). In addition, the results of this study provide an initial basis for the potential of eye-tracking as a diagnostic tool for detecting student conceptions and difficulties in distributional comparison.

Addressing our second research question, we analyzed potential differences between 6th graders and 8th graders in terms of viewing behavior and statistical thinking. The results of the stimulated recall interviews revealed that half or more of the 6th as well as the 8th graders considered global features either in none of the items or in all four items. Thus, both samples have in common that students showed a certain consistency in their statistical thinking across all items. No significant difference was found between 6th and 8th graders in terms of the relative amount of items in which at least one global feature was perceived and processed. Likewise, the analysis of the eye-tracking measures did not reveal any significant differences between both groups. Essential for statistical data analysis is that it is mainly about describing global features of data distributions (e. g., Bakker & Gravemeijer, 2004; Ben-Zvi, 2004). Against this background, the students who did not or only rarely considered any global characteristics across all items are to be classified as problematic. These students remained with local strategies (e.g., comparing absolute frequencies of dots in certain intervals), even if sample sizes are unequal, which is an incorrect strategy in these cases. These findings are consistent with existing research describing students' difficulties in understanding a distribution as a whole, which appear to persist even after instruction in statistics (e. g., Ben-Zvi & Arcavi 2001; Watson & Moritz, 1999). Although 8th graders, according to the curriculum, should be more advanced in interpreting data distributions (e.g., in terms of spread/density), there were no significant differences between 6th and 8th graders in our sample, neither in their gaze behavior nor in their statistical thinking in terms of perceiving and processing global distributional features.

This study is a first step towards enhancing our understanding of students' visual attention and associated statistical thinking when comparing data distributions. Future research is necessary to examine potential influencing factors on the part of students (e.g., topic specific pre-knowledge) and performance differences within the groups of local and global viewers in more detail. In addition, the results of this study may also

build a starting point for future research investigating the potential of eye-tracking as a diagnostic tool that can be used in teacher training or school practice to detect and learn about students' conceptions and difficulties in distributional comparison.

REFERENCES

Bakker, A., & Gravemeijer, K. P. E. (2004). Learning to Reason About Distribution. In D. Ben-Zvi & J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning and Thinking* (pp. 147–168). Springer Netherlands.

Ben-Zvi, D. (2004). Reasoning about variability in comparing distributions. *Statistics Education Research Journal*, *3*(2), 42–63.

Ben-Zvi, D., & Arcavi, A. (2001). Junior high school students' construction of global views of data and data representations. Educational Studies in Mathematics, 45(1), 35–65. https://doi.org/10.1023/A:1013809201228

Boels, L., Bakker, A., & Drijvers, P. (2019). Eye tracking secondary school students' strategies when interpreting statistical graphs. *Proceedings of the 43rd Conference of the International Group for the Psychology of Mathematics Education*. PME.

Frischemeier, D. (2019). Primary school students' reasoning when comparing groups using modal clumps, medians, and hatplots. *Mathematics Education Research Journal*, *31*(4), 485–505.

Gal, I., Rothschild, K., & Wagner, D. A. (1989). Which Group Is Better? The Development of Statistical Reasoning in Elementary School Children. *Meeting of the Society for Research in Child Development* (Kansas City, MO, April 1989).

Holmqvist, K., & Andersson, R. (2017). *Eye tracking: A comprehensive guide to methods, paradigms, and measures* (2nd edition). Lund Eye-Tracking Research Institute.

Konold, C., & Higgins, T. L. (2003). Reasoning about Data. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 193–215). National Council of Teachers of Mathematics.

Konold, C., & Pollatsek, A. (2004). Conceptualizing an Average as a Stable Feature in a Noisy Process. In D. Ben-Zvi & J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning and Thinking* (pp. 169–199). Springer Netherlands.

Schindler, M., & Lilienthal, A. J. (2019). Domain-specific interpretation of eye tracking data: towards a refined use of the eye-mind hypothesis for the field of geometry. *Educational Studies in Mathematics*, *101*(1), 123–139.

Schreiter & Vogel (in press). Eye-tracking measures as indicators for a local vs. global view of data [Special Issue]. *Frontiers in Education*.

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*(2), 147–200.

Watson, J. M., & Moritz, J. B. (1999). The Beginning of Statistical Inference: Comparing Two Data Sets. *Educational Studies in Mathematics*, *37*(2), 145–168.

HOW DO MATHEMATICS TEACHERS LEARN TO CREATE A MATHEMATICAL STORYLINE IN PROBLEM-BASED LESSONS?

Gil Schwarts, Patricio Herbst and Amanda Brown

University of Michigan, USA

*Building on student work (SW) in mathematics classroom discussion requires complex decision-making from mathematics teachers. Previous literature on problem-based lessons recommends selecting and sequencing pieces of SW in a way that creates a mathematical storyline, but there is rarely any empirical evidence on how mathematics teachers can master such practices. We use the case of Story*Circles, *a lesson-based professional development program, to show how iterative processes in which teachers were engaging with SW assisted them in developing heuristics for a careful selection and sequencing of SW. The results show that these processes involved 1) the teachers' emerging awareness of features of SW; and 2) an evolving capacity to relate these features to the lesson goal. We discuss design features that fostered these changes.*

BACKGROUND AND RATIONALE

Building on student thinking in mathematics lessons is a core aspect of responsive teaching. Instructional practices purported to support teachers' attempts to foreground student ideas include posing mathematically rich problems, monitoring students as they work on such problems, and then selecting several students to share work in a sequence that can be leveraged to support a productive whole-class discussion. The latter two practices are known as *selecting* and *sequencing* (Smith & Stein, 2011). The literature recommends that teachers select and sequence in ways that create "mathematically coherent storylines" (ibid, p. 44). Like any other good story, a crucial component of a mathematical storyline is its culmination, which, in the case of mathematics lessons, means that the discussion of a problem leads to the mathematical goal of the lesson (Kazemi & Hintz, 2014). Despite the importance of the lesson goal, empirical studies show that teachers tend to overlook it when justifying how they select and sequence (Ayalon & Rubel, 2022; Dunning, 2022). This suggests that attending to the lesson goal when selecting SW is part of teachers' tacit knowledge (Herbst & Chazan, 2011). For these reasons, we suspect that teachers' learning of such practices may be enhanced through participation in "infrastructures that support the interplay of knowledge and knowing" (Cook & Brown, 1999, p. 381). Below, we describe how a collaborative practice-based professional development (PD), Story*Circles* (details below), supported teachers in sharing and expanding their ways of knowing related to a careful selection of SW. Compared to the abundance of literature about classroom discussions, the practices of selecting and sequencing are under-researched, despite their importance. Our contribution to this burgeoning body of research is in unpacking the notion of "mathematical storylines" by exploring how geometry teachers became explicit and deliberate about their decisions related to SW. We ask:

2023. 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. 187-202). PME 46.

*What justifications for selecting and sequencing SW do teachers make available when participating in Story*Circles*? How do teachers' justifications evolve over time?*

THEORETICAL FRAMING

To explore how teachers make decisions about SW we draw on some of the key ideas of the *practical rationality* framework (Herbst & Chazan, 2011). First, we build on the idea that teachers' actions are always related to the norms of the instructional situation in which they operate. Instructional situations are events in classrooms that teachers recognize as familiar and mentally categorize as a situation of a particular type (Herbst & Chazan, 2011). In the US high-school context, common instructional situations in geometry include *construction* and *proof* (Herbst et al., 2018). It follows that the norms which are related to particular instructional situations are subject-specific, that is, allude to the mathematical content that is being taught. Second, we apply this idea to describe categories of teacher perception (Herbst et al., 2021), namely, aspects to which mathematics teachers attend when they examine and make decisions upon SW: the category of normativity alludes to the SW alignment with the teacher expectations, based on how the teacher framed the instructional situation. For example, if a problem is framed as one expecting students to do a construction, teachers may consider a sketched figure as less normative than a SW where construction tools were used. A second category is the serviceability of the SW, which attends to the alignment with the lesson goal. That is, a sketch that suggested a connection to the lesson goal might be considered more serviceable (even though it is less normative) than a construction that provides no leverage for progress. This emerging framework enables us to discern subject-specific aspects of teachers' decisions.

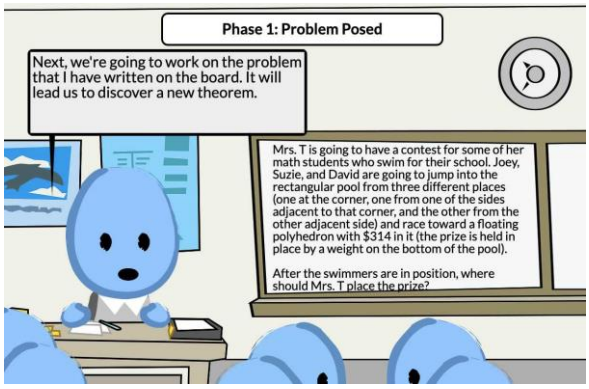
METHODS

The design of the PD environment

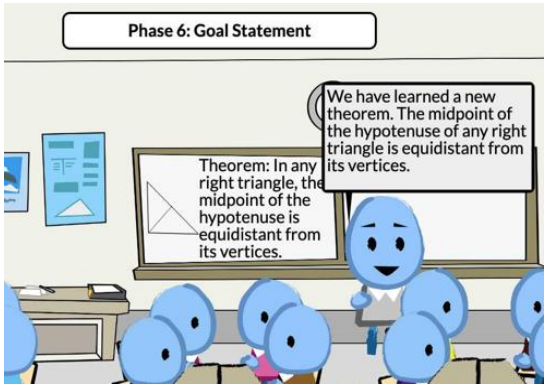
Story*Circles* is a collaborative PD where secondary mathematics teachers anticipate a lesson through iterative phases of scripting, visualizing, and arguing about it (Herbst & Milewski, 2018), in online synchronous and asynchronous activities. Inspired by Japanese lesson study, each Story*Circle* focuses on participants' attempts to improve one lesson, initially sketched in a storyboard, as they see fit. Importantly, the goal of Story*Circles* is to foster teachers' peer argumentation about practice (and not, for example, to direct them to teach a specific lesson or to include specific moves).

At the beginning of a Story*Circle*, participants view one version of the storyboarded lesson. Over six weeks they are engaged in various activities where they discuss key decision points in the lesson and script more scenes accordingly. The constants of the lesson-to-be-revised are the posed problem (the first frame of the storyboarded lesson, see Figure 1a) and the culminating institutionalization of the instructional goal (last frame, Figure 1b). In the particular Story*Circle* considered here, the pool problem lesson (in geometry) was represented with nine frames, that showed the arc of the lesson as a sequence of phases: Problem Posed (Phase 1); Getting Your Feet Wet (Phase 2); Whole Class Check-In (Phase 3); Redirecting the Work (Phase 4); Whole

Class Discussion (Phase 5); Goal Statement (Phase 6). The lesson starts with a problem that US teachers tend to frame as a construction (Figure 1a) and culminates with arriving at the theorem stating that "the midpoint of the hypotenuse of any right triangle is equidistant from its vertices" (See Figure 1b). A main resource for the participants' work on the lesson is a collective examination of samples of SW, as will be further detailed. The facilitator was an experienced geometry teacher who previously participated in Story*Circles*. A diverse group (in terms of gender, ethnicity, experience, institutions, and more) of seven teachers participated in the entire PD cycle.



1a



1b

Figure 1. Scenes at the beginning and the end of the pool problem lesson (© 2021, The Regents of the University of Michigan, used with permission)

Data sources and analysis

Focusing on participants' decision-making when interacting with SW, the data sources used for this analysis are videos and transcripts from two synchronous meetings in which SW was the focus: the second meeting (M2) and the fifth (M5). To identify how participants' justifications were made available to peers through discussions, we performed a content analysis guided by the categories of normativity and serviceability (Herbst et al., 2021). The code "normativity" was used when participants referred to the norms of the instructional situation of construction (e.g., when mentioning tools or precision) to justify a decision to select or sequence SW in particular ways. The code "serviceability" was used when they referred to the lesson goal (mediated by mentioning mathematical ideas or objects) to justify a decision to select or sequence SW. The process included: (a) segmenting the transcript into sets of utterances associated with each SW sample and then into idea units that include justifications; (b) coding the segments using the top-down categories together with bottom-up themes; (c) comparing the analyses of M2 and M5, focusing on the place of the instructional goal – namely, attention to serviceability – in participants' justifications; (d) identifying key moments in the talk that illustrate the evolution of justifications.

FINDINGS

The analysis identified that participants' focus on serviceability shifted across their participation in Story*Circles*, with them hardly attending to serviceability in M2 (30% of codes) while serviceability became a core aspect of their arguments in M5 (74% of codes). The following processes were identified: (1) The group's emerging awareness of serviceable aspects of SW; and (2) an evolving capacity to relate these aspects to the lesson goal, while disregarding other aspects. Below, we illustrate these processes. "F" is short for facilitator, and all of the participants' names are pseudonyms.

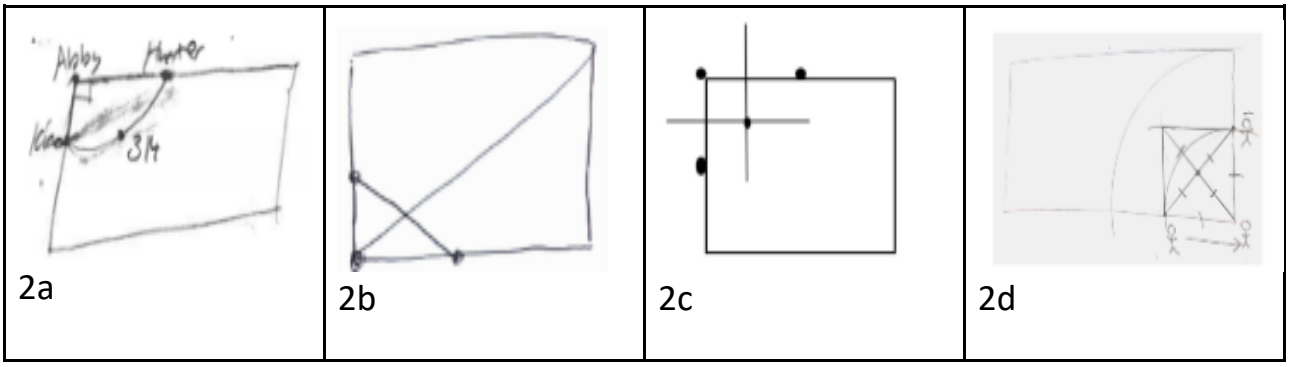


Figure 2. Samples of student work that were discussed in M2 (© 2021, The Regents of the University of Michigan, used with permission)

WEEK 2 MEETING (M2): OVERLOOKING THE LESSON GOAL

Prior to M2, the teachers participated in an introduction session and two asynchronous activities in which they perused and annotated an initial version of the pool problem lesson and nine pieces of SW, respectively (see examples for the SW in Figure 2). M2, which was aimed at focusing participants' attention on the practices of selecting and sequencing SW, began with the question: "What pieces of work catch your attention? [...] Are there ideas that you particularly want to make sure that you brought out?". The participants mainly attended to the first question, discussing SW that were unexpected (Figure 2a) or non-normative (Figure 2b), and then coalesced toward selecting Figure 2b as the first work they want to be shared on the board:

246 Ira: There'd be more than one person who just draws the rectangle, puts the swimmers there, [I'd] say okay where are your tools? What did you do?

273 Ran: I'd rather start with this [Figure 2b], because I think everyone would understand [...] I think it's better to start out with a more general one.

We hear the participants as claiming that selecting 2b would be justifiable on account of their perceiving that piece of work as accessibile, general, representative, and non-normative – the latter aspect suggests that the SW was partly selected as a non-example that is used to direct students to follow the norms of the situation (using construction tools). Notably, none of the justifications mentioned the lesson goal. The facilitator then asked which SW could be useful, a move that led to the first emergence of serviceability in a justification; one participant suggested Figure 2c, "Obviously,

because that's a perpendicular bisector, that gets right to where you're trying to go." The word "obviously" suggests that justification for selecting SW that aligns with the lesson goal seems to her as taken-as-shared knowledge that might not be worth mentioning. Following the failed attempts to discuss the lesson goal based on features of the SW, the facilitator tried again:

341 F: What would be the mathematical ideas that we want to make sure got out on the table? Can you imagine bringing up SW to pull these ideas out?

Here, the facilitator illustrates a heuristic for a mindful selection of SW: First thinking about the mathematical ideas that are needed for the discussion, and then selecting SW that includes these ideas or that can be used to prompt discussions about them. Nevertheless, participants mostly suggested ideas that they noticed in the SW such as Mentions "equidistant"; Properties of a rectangle; Right triangle; and Circle, overlooking the hypotenuse, midpoint, and perpendicular bisectors. As the discussion evolved, participants gradually attended to serviceable aspects (such as the generality of the theorem). Yet, the facilitator noticed that they were still not prioritizing any SW, wishing to be attentive to each and every mathematical idea and SW.

454 F: Are these all equally important ideas or do we want to make sure we're prioritizing some of these over others? The goal is to be able to have the students have a discussion that ends up with them discovering this theorem.

458 Clader: It kind of already is in order [...] talking about the equidistant part first, recognizing the rectangle […]. The only thing that obviously would be a little thrown off is if you were going to do the circles piece, I think that might not naturally lead into it, but maybe too [...]

This exchange shows the emergence of the understanding that presenting all pieces of SW, and discussing all related ideas, could impede the coherence of the lesson ("a little thrown off"). A few minutes later, the group made further progress in this direction:

496 Clader: What we need to decide is, like, which direction we're going, because there are so many different entry points to this […] Now we have all these ideas, and that's kind of where I'm stuck because […] once we decided a way [...] I think it'd be easier to determine what drawing next I'd want to lift.

501 Ira: Yeah that's where I'm stuck too [...] what if someone use the tools that found the midpoint and found the right answer and have no clue how they got it, but they're putting up the correct drawing?

This exchange shows the processes in which the participants realize the need to prioritize, an insight that emerged together with the understanding that there are multiple storylines that could be developed and focusing on one of them requires making decisions. The word "stuck" suggests that the participants' conflict is not only about what decision to take, but also about identifying that this moment involves complex decision-making that takes account not only of the features of a particular piece of SW, but also of how those features serve the goal of the lesson. That leads to an evolving awareness of more features of SW (Turn 501), where the participant claims that not all of the SW which is correct and/or normative is also useful for the whole class discussion. Overall, we identified two constraints for the creation of a

mathematical storyline: 1) adherence to normativity; 2) a desire for full attentiveness, including making sure that all givens are used, instead of focusing on the question.

WEEK 5 MEETING (M5): ATTENDING TO SERVICEABILITY

M5 was conducted after the group further engaged with the samples of SW in several asynchronous activities which included, for example, scripting classroom dialogues. The meeting described below focused on the last two phases of the lesson, where the second whole class discussion (Phase 5) leads to the discovery of the theorem (Phase 6). The meeting began with the facilitator asking "What would you like to have on the table going into that last class discussion?". This time, the conversation was ample with arguments mentioning the lesson goal:

267 Ran: Most of what we had [in previous phases] had to do with rectangles, and nothing to do with hypotenuse [...] , and even the radii and the arches, those are ideas to measure, maybe, but not necessarily **conducive to the conclusion that we're looking for**.

The facilitator then asked what kind of work the participants imagine can help generate this conversation. That led to a discussion about the normativity of the work. One participant, instead of imagining, attended to his reality, saying: "I think students would draw it [the pool] rather than construct it, unless you told them to construct it" [272]. The facilitator then asked if the group thought it was important that the students construct. While in M2 participants had expressed their expectation for constructions without providing an explanation why, in this meeting most of their comments implied that construction is important for the sake of achieving the lesson goal. In other words, they saw construction as a valid alternative to proof for arriving at the theorem. The following is an illustrative exchange:

303 Clader: Towards the very end I would like to see something more formal.

309 Ran: Without doing a construction, there has to be a [different] way, because otherwise we're just theorizing. How do we know that it's equidistant? [...] At least my students, [...] if I say it's equidistant they will agree, but that does not prove. It's conducive to them to conclude that on their own.

315 Llara: I agree, I think […] that would be where you can now show them why a construction is used instead of just answering question 23 on the test. So they can actually see that the construction now will verify.

324 Ran: And I'm thinking that what we say at the class discussion, I think there has to be something in their work that is what we're going to discuss about.

338 Labrona: I would kind of hope [that] some of the work included a perpendicular bisector so that might be something we can carry into the story, **because for me this is the justification for the eventual goal about the midpoint of the hypotenuse**.[...] I think there were a couple pieces of work that either alluded to a perpendicular bisector or more explicitly had one.

The group then selected Figure 2c and discussed how it could be further used in a discussion. Then, they returned to the issue of which student solutions are considered valid for them. In this discussion, they were not adhering to normativity anymore:

494 F: I wonder if you've got feelings about the idea of a formal or informal proof, or what would make for just a valid conversation that would satisfy?

496 Ran: I like the idea of allowing them to explore different avenues, so I'm not set on formal or informal. […] if one wants to write a proof and the other one wants to use a compass, then that's okay, **as long as they do learn the concept that I want them to learn**.

501 Llara: At least for me, as long as the students can explain it, that's the most important thing, I mean it doesn't have to be formal or informal, [...] what's important [is] that it's correct in the way they explained it for us.

532 Clader: we're trying to emphasize specific vocabulary to get them to the end [...] **the pieces of vocabulary just have to be threaded throughout each example to get them to our final conclusion**. I don't think it's a problem of how they got there as long as we can emphasize those specific pieces.

This exchange illustrates awareness of more features of SW, as well as an increased ability to set aside the procedures to solve the problem and emphasize, instead, the achievement of the instructional goal. In the participants' justifications, there is an emerging recognition of the importance of emphasizing the serviceable aspects of SW.

Overall, in this session teachers could better justify why they selected or disregarded SW, by recognizing features that allude to the lesson goal. Consequently, their choices created a blueprint for a coherent mathematical storyline.

DISCUSSION

This paper examines how teachers' justifications for selecting and sequencing were made available and evolved during their participation in Story*Circles*. We showed how teachers' decision-making became more explicit and deliberate, that is, related to the lesson goal (as Kazemi & Hintz, 2014; Smith & Stein, 2011 recommend). We highlight two notable findings: 1) by noticing more features in SW teachers were able to better justify their choices; and 2) teachers adopted the idea that creating a mathematical story involves prioritizing work that attends to the lesson goal. Although such prioritization could be in tension with attentiveness to all students (Ayalon & Rubel, 2022), teachers reconciled this tension by letting the lesson goal guide their decisions.

The following elements contributed to the emerging focus on serviceability: 1) **The phase structure of the storyboarded lesson**. Based on previous iterations of Story*Circles*, the design of the current cycle provided teachers with the arc of the lesson – that is, a structure that conveyed the sense of what is happening in each phase and how particular moments are related to the larger context of the lesson. This feature enables zooming in and out and relating decisions into larger goals; 2) **Iterative engagement with pieces of SW**. The teachers had diverse opportunities to engage with the SW, with peers and alone, in iterative activities in which the lesson goal has gradually become more salient in their interpretations. We maintain that developing heuristics for purposeful decision-making in an environment of reduced complexity is essential for a later adoption of such heuristics; 3) **Responsive facilitation**. The facilitator was attentive and hardly provided input, yet she was deliberate on building

a coherent storyline in her navigation of teachers' talk. Moreover, she was modeling how to lead a discussion that builds on learners' ideas and is goal-oriented.

The contribution of this work is in unpacking how teachers can be supported in creating a mathematical storyline (Smith & Stein, 2011), by using a subject-specific lens that is sensitive to the mathematics at the core of teachers' decision-making.

Acknowledgements. This work has been done with the support of Grant 220020524 from the James S. McDonnell Foundation. All opinions are those of the authors and do not represent the views of the Foundation.

REFERENCES

Ayalon, M., & Rubel, L. H. (2022). Selecting and sequencing for a whole-class discussion: Teachers' considerations. *The Journal of Mathematical Behavior*, *66*, 100958. https://doi.org/10.1016/j.jmathb.2022.100958

Cook, S. D., & Brown, J. S. (1999). Bridging epistemologies: The generative dance between organizational knowledge and organizational knowing. *Organization science*, *10*(4), 381-400.

Dunning, A. (2022). A framework for selecting strategies for whole-class discussions. *Journal of Mathematics Teacher Education*. In press. https://doi.org/10.1007/s10857-022-09536-5.

Herbst, P. & Chazan, D. (2011). Research on practical rationality: Studying the justification of actions in mathematics teaching. *The Mathematics Enthusiast*, *8*(3), 405-462.

Herbst, P., Boileau, N., & Gürsel, U. (2018). Examining the work of teaching geometry as a subject specific phenomenon. In P. Herbst, U. Cheah, K. Jones, & P. Richard (Eds.), *International Perspectives on the Teaching and Learning of Geometry in Secondary Schools* (pp. 87-110). Springer.

Herbst, P., Chazan, D., & Boileau, N. (2021). Framing, normativity, and serviceability in teachers' decision making during lessons. Paper accepted for presentation at TSG 37, ICME 14, Shanghai, China, 2021, later withdrawn. Available at https://hdl.handle.net/2027.42/171072

Herbst, P. & Milewski, A. M. (2018). What StoryCircles can do for mathematics teaching and teacher education. In R. Zazkis, & P. Herbst (Eds.), *Scripting approaches in mathematics education: Mathematical dialogues in research and practice* (pp. 321–364). Springer.

Kazemi, E., & Hintz, A. (2014). *Intentional talk: How to structure and lead productive mathematical discussions.* Stenhouse.

Smith, M., & Stein, M. (2011). *Five practices for orchestrating productive mathematical discussions*. National Council of Teachers of Mathematics.

MATHEMATICAL REASONING TYPES AS GENDERED? VIEWS FROM PALESTINIAN/ARAB ISRAELI TEACHERS

Juhaina Awawdeh Shahbari[1], Laurie Rubel[2] and Fatema Kabha[1]

[1]Al-Qasemi Academic College, [2] University of Haifa

The study examines how mathematical reasoning is socially gendered among Palestinian/Arab Israeli teachers. Middle and high school mathematics teachers (n=49) in Arabic-language schools participated by first classifying five solutions to each of five mathematics questions according to: procedural (school algorithm), creative-logical (non-standard), or guess-and-check. Participants then were prompted to attribute each solution to four provided categories: high/low achieving girl or boy. Participants attributed most (68%) of the creative/logical solutions to high-achieving boys, most (69%) of the procedural solutions to high-achieving girls, and most (66%) of the guess-and-check solutions to low-achieving boys.

INTRODUCTION

We conducted this study within the social context of Palestinian/Arab-Israelis (P/AIs), who represent about 21% of Israel's citizenry.[1] There are contradictions related to gender and mathematics education in this context. P/AI women are marginalized–as women, as Arabic speakers, as Israeli citizens– and the various intersections of these dimensions. This marginalization is evident in how, for example, in 2018, only 50% of P/AI women between the ages of 18-22, were employed or in school (Haj-Yahya et al., 2018). At the same time, in Israel's Arabic-language schools, girls tend to outperform boys–on state and international mathematics tests and at all school levels– and are better represented in advanced mathematics, physics, and computer science coursework (Pinson et al., 2020). This latter trend has attracted research attention by challenging the dominant pattern in Israel of a gender gap in mathematics favoring boys, with questions as to whether and how mathematics has been socially constructed in the P/AI context (e.g., Forgasz & Mittelberg, 2008).

School success of P/AI girls and women is socially constrained and has not yet translated for girls and women into participation in related professional (and more lucrative) fields. Instead, even though there is a surplus of teachers in Arabic-language schools, P/AI girls and women who excel in mathematics at the secondary and university levels tend to become mathematics teachers (Fuchs & Wilson, 2018). Gender norms around childcare, familial pressures, and the conduciveness of working in schools to remaining within P/AI communities are influential (Rubel & Ehrenfeld, 2020). However, ways in which success in mathematics is socially gendered could be another factor that explains why the majority of high-achieving P/AI women become

[1] Categorized formally Israel as "Arab Israeli," the authors instead opt for the term Palestinian/Arab Israeli. This acknowledges the local identifier and accounts for increasingly prevalent public expressions of Palestinian identity and solidarity.

2023. 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. 195-202). PME 46.

teachers rather than continuing to mathematics-related professions. Prior research has documented gendered narratives in circulation in the United States and Europe, for example, in which boys who are successful in mathematics are cast as brilliant, in contrast with girls as diligent (Heyder et al., 2019). The social construction of success in mathematics as being shaped by diligence rather than brilliance, but only for women, acts as an obstacle, in terms of identity formation and how they view their own success as well as how it is viewed by others. The current study explores the gendering of mathematics in terms of various types of mathematical reasoning.

PRIOR RESEARCH

Despite social advances towards gender equity, gender-related biases continue among teachers with respect to a range of everyday classroom practices, around the world. Studies in the United States and Ireland, for example, show teachers' estimation of boys' mathematics ability as higher than girls, relative to the same mathematical work (Copur-Gencturk et al., 2020; McCoy et al., 2022). In Lebanon, teachers were shown to tend to attribute boys' success to ability but girls' success to effort (Sarouphim & Chartouny, 2017). In Israel, by comparing classroom mathematics grades with standardized national exam scores, by student, teachers tended to give boys higher grades relative to the same level of work (Lavy & Sand, 2018).

The findings of these studies collectively reflect social narratives that connect ability in mathematics, writ large, to maleness. Others have looked more specifically at the social gendering of mathematics in terms of specific mathematical reasoning types. We can distinguish between creative and imitative mathematical reasoning (Lithner, 2008): creative reasoning is marked by novelty, plausibility, and mathematical logic, whereas imitative reasoning consists of memorization or implementation of algorithms. Creative reasoning is seen to be indicative of mathematical smartness, while imitative reasoning is considered to be a form of compliance or rule following.

At the outset, it is important to note that across time and contexts, there is a pattern of girls being more likely than boys to use concrete, familiar strategies (Cimpian et al., 2016). We view gender as socially constructed, and our's is not a study of biological tendencies with respect to mathematical reasoning. Rather, we understand such a pattern to be a product of socialization processes, wherein teachers are important mediators. Sumpter (2016), for example, found that teachers in Sweden tend to attribute guessing to boys and algorithmic reasoning to girls, with unclear results about gendering of creative/logical reasoning, which they had anticipated might be attributed to boys. In Authors' previous study, when P/AI teachers were prompted as to whom they would select to present their mathematics work to the class, they mostly selected a girl to present a direct model solution and a boy to present a creative solution. The instrument from the previous study included only one mathematical task, prompting the need for a follow-up, broader study. Here we pursue this research question: In what ways do teachers relate different types of mathematical reasoning to gender and why?

METHODS

We recruited participants by distributing an online questionnaire using contact groups of teachers, summing to 49 in-service P/AI mathematics teachers (Table 1). In a second phase, we interviewed 10 participants (eight women and two men).

Table 1. Participants' background characteristics

Background Characteristics		Women	Men	Total
Gender		32(65%)	17(35%)	49(100%)
School Level	Middle (Grades 7-9)	14(67%)	7(33%)	21(43%)
	High (Grades 10-12)	18(64%)	10(36%)	28(57%)

Data Sources, Procedures and Analysis

We designed an instrument in Arabic consisting of five mathematics questions that lend themselves to be solved using various strategies. For example, one problem read: A farm has chickens and cows. All together, among them, there are 70 heads and 186 feet. How many chickens and how many cows are on the farm? (Tabach & Friedlander, 2013). For each question, we presented five different solutions (examples in Table 2). A team of experts (consisting of the authors and four other experienced mathematics teachers) classified the solutions according to: procedural, creative or logical, and guess-and-check. The procedural solutions are standard, school-taught algorithms. Creative solutions are those that use novel or unconventional strategies, oftentimes relying on logical reasoning. Guess-and-check reasoning indicates an educated guess about a possible solution with iterative tests and improvements to each guess.

Table 2. Examples of solutions to Chicken/Cow problem (translated to English)

Procedural	Creative or Logical	Guess-and-Check
X= number of cows y=number of chickens 2x= legs of cows 4y=legs of chickens $x + y = 70$ $4x + 2y = 186$ ——— $2x + 2y = 140$ $4x + 2y = 186$ ——— $-2x = -46$ $x = 23$ cows $23 + y = 70$ $y = 47$ chickens	Group= 3 animals cow= 2 chickens $70 : 3 = 23\frac{1}{3}$ We build 23 groups, and each group contains a cow and two chickens, so we have 23 cows and 46 chickens, and the number of legs is: $23 \cdot 4 + 46 \cdot 2 = 92 + 92 = 184$ $186 - 184 = 2$ The last animal is a chicken and has two legs Cows=23 Chickens= 46+1=47	Suppose the number of cows is equal to the number of chickens, 35 of each. The number of legs will be: 35·2+35·4=70+140=210 Since the answer is more than 186, we should reduce the number of cows and increase the number of chickens until we get to 186. | # chickens | # of cows | # of legs | | 35 | 35 | 210 | | 36 | 34 | 208 | | 47 | 23 | 186 |

We prompted participants to classify each solution according to categories: school procedure, creative or logical, or guess-and-check. In total, there were 392 (8 x 49) procedural solutions and responses, 539 (11 x 49) creative-logical solutions and responses, and 294 (6x49) guess-and-check solutions and responses.

We prompted participants to select whom they assess has solved the problem in each way, from among the four categories: high-achieving boy/girl or low-achieving boy/girl. We are aware that in classifying people as boys or girls, our instrument reinforces a false gender binary, yet this is a social order that carries meaning and practicality in the local context (see Hall, 2014). The instrument did not make available a response of "there is no way to tell who solved the problem in this way," which possibly would have been favored by some participants.

In a second phase, we interviewed ten participants using a structured protocol in which we presented the quantitative results from the group of participants and asked each interviewee to interpret the results. We recorded and transcribed the interviews and used qualitative thematic analysis (Braun & Clarke, 2006).

FINDINGS

Classification of solutions

Participants' classifications mostly matched the expert team's (75% of the procedural, 84% of the creative/logical, and 77% of the guess-and-check). We discarded all instances in which a participant's classification did not match the expert team's. We applied this matching as a filter as a way to be sure that participants viewed a particular solution as indicative of the corresponding form of reasoning. This filtering resulted in 974 (80%) cases. The 974 resultant cases include 294 procedural, 455 creative/logical, and 225 guess-and-check solutions.

Participants' attributions of solutions

Overall, participants attributed solutions to boys (60%) more often than girls (40%). Participants attributed creative solutions and guess-and-check solutions much more often to boys and standard procedural solutions much more often to girls (Table 3). A chi-squared test shows that the association between mathematical reasoning types and attribution by gender is statistically significant, χ^2 (4, N = 974) =192, p < .001.

Table 3. Attributions of solutions by gender

	Girls	Boys	Total
Procedural (standard)	215(73%)	79(27%)	294
Creative or Logical (non-standard)	110(24%)	345(76%)	455
Guess-and-Check	66(30%)	159(70%)	225
Total	391 (40%)	583(60%)	974(100%)

Participants attributed solutions to each of the four student categories but not uniformly so (Table 4). The high-achieving categories (boys or girls) were selected most often (70%). High-achieving boys were noted with the highest frequency, and most (68%) of the creative or logical solutions were attributed to them. In contrast, most (69%) of the procedural solutions were attributed to high-achieving girls. Finally, most (66%) of the guess-and-check solutions were attributed to low-achieving boys. Overall, low-achieving girls were selected with the lowest frequency, receiving only 9% of all attributions. A chi-square test shows that the relationship between solution type and student achievement level with gender is statistically significant χ^2(6, N = 974) =782.34, p < .001.

Table 4. Attributions of solutions by gender and achievement

	High-achieving		Low-achieving		
	Girl	Boy	Girl	Boy	Total
Procedural	203(69%)	57 (19%)	12(4%)	22(8%)	294
Creative / Logical	96(21%)	307 (68%)	14(3%)	38(8%)	455
Guess-and-check	5(2%)	11(5%)	61(27%)	148(66%)	225
Total	304(31%)	375 (39%)	87(9%)	207(21%)	974

Explanations

In general, interviewees were not surprised by the summary frequencies of the attributions by achievement and gender. Most (seven) explained these trends by retelling essentialist narratives, such as “Boys are exposed to more things than girls. That’s why boys are more creative” or ”girls prefer to follow step-by-step procedures even if they are laborious” or “girls are hesitant to guess because they fear getting a wrong answer.” One interviewee related these results to an achievement gap that favors girls: “In class, boys demonstrate greater creativity. But on exams, girls are more successful because of their ability to solve questions using taught strategies.“ Three interviewees challenged the findings in various ways: one said that these patterns vary by mathematics topic; another questioned that the group attributed guess-and-check specifically to boys. One interviewee emphasized that she thinks that gendered reasoning types are cultivated by teachers and not innate to any group of students.

DISCUSSION

Participants tended to attribute creative or logical solutions to high-achieving boys, procedural solutions to high-achieving girls, and guess-and-check solutions to low-achieving boys. Explanations in the interviews revealed gendered narratives about mathematical reasoning that explain these patterns. The patterns of attributions suggest that P/AI teachers tend to view boys as more capable of independent thought and able to invent their own strategies, but view girls as more likely to apply previously learned algorithms. Guess-and-check solutions were attributed most often to low-achieving boys, in correspondence with views of guessing as risk-taking and different from “safer” forms of procedural reasoning. Low-achieving girls were the most underrepresented, consistent with previous research in other contexts (Jones, 2005). Gender gaps in mathematics among P/AIs (and across the Middle East) persistently favor girls, different from most of the world. Previous studies have suggested that one possible interpretation is that the social gendering of mathematics is different in this region (e.g., Forgasz & Mittelberg, 2008). Our findings, however, align with those of previous studies in other contexts (e.g., Copur-Gencturk et al., 2020; McCoy et al., 2020; Sarouphim & Chartouny, 2017). Social narratives about mathematical reasoning as gendered like these do not help us to understand the relative success of girls in school mathematics, but shed light on the later under-representation of P/AI girls in mathematics-related professions. School success in mathematics might possibly

reinforce or even exacerbate the positioning of girls as holding imitative, rather than creative, reasoning abilities, meaning that their success at the school level is being undercut by this and other social processes.

This research was supported by the Israel Science Foundation (grant No. 2747/21).

REFERENCES

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101.

Cimpian, J. R., Lubienski, S. T., Timmer, J. D., Makowski, M. B., & Miller, E. K. (2016). Have gender gaps in math closed? Achievement, teacher perceptions, and learning behaviors across two ECLS-K cohorts. *AERA Open, 2*(4), 1–19.

Copur-Gencturk, Y., Thacker, I., & Quinn, D. (2020). K-8 teachers' overall and gender-specific beliefs about mathematical aptitude. *International Journal of Science and Mathematics Education, 19,* 1251-1269.

Forgasz, H., & Mittelberg, D. (2008). Israeli Jewish and Arab students' gendering of mathematics. *ZDM Mathematics Education, 40*(4), 545–558.

Fuchs, H., & Wilson, T. F. (2018). Arab-Israeli women entering the labor market. Report for the Taub Center for Social Policy Studies in Israel http://taubcenter.org.il/wpcontent/files_mf/arabisraeliwomeninthelabormarket.pdf/

Haj-Yahya, N. H., Schnell, I. & Khattab, N. (2018). The exclusion of young Arab women from work, education and training in Israel. *Quality & Quantity: International Journal of Methodology, 52*(1), 157-173.

Hall, J. (2014). Unpacking "gender issues" research. *Philosophy of Mathematics Education Journal, 28,* 1–10.

Heyder, A., Steinmayr, R., & Kessels, U. (2019). Do teachers' beliefs about math aptitude and brilliance explain gender differences in children's math ability self-concept? *Frontiers in Education, 4,* 1–11.

Jones, S. (2005). The invisibility of the underachieving girl. *International Journal of Inclusive Education, 9*(3), 269–286.

Lavy, V., & Sand, E. (2018). On the origins of gender gaps in human capital: Short-and long-term consequences of teachers' biases. *Journal of Public Economics, 167,* 263-279.

Lithner, J. (2008). A research framework for creative and imitative reasoning. *Educational Studies in Mathematics, 67*(3), 255–276.

McCoy, S., Byrne, D., & O'Connor, P. (2022). Gender stereotyping in mothers' and teachers' perceptions of boys' and girls' mathematics performance in Ireland. *Oxford Review of Education, 48*(3), 341-363

Pinson, H., Feniger, Y., & Barak, Y. (2020). Explaining a reverse gender gap in advanced physics and computer science course-taking: An exploratory case study comparing Hebrew-speaking and Arabic-speaking high schools in Israel. *Journal of Research in Science Teaching, 57*(8), 1177–1198.

Rubel, L., Ayalon, M., Shahbari. J. A. (2022). Sequencing and selecting solutions in a gendered world. *Mathematical Thinking and Learning*, 1-23. DOI: 10.1080/10986065.2022.2062641

Rubel, L., & Ehrenfeld, N. (2020). Palestinian/Arab Israeli women's experiences in mathematics education: an intersectional analysis. *International Journal of Educational Research*, 102.

Sarouphim, K. M., & Chartouny, M. (2017). Mathematics education in Lebanon: Gender differences in attitudes and achievement. *Educational Studies in Mathematics, 94*(1), 55-68.

Sumpter, L. (2016). Investigating upper secondary school teachers' conceptions: Is mathematical reasoning considered gendered? I*nternational Journal of Science and Mathematics Education, 14*(2), 347-362.

Tabach, M., & Friedlander, A. (2013). School mathematics and creativity at the elementary and middle-grade levels: how are they related? *ZDM Mathematics Education, 45*(2), 227-238.

DYNAMIC VISUALIZATION AND EMBODIED DESIGN FOR TRIGONOMETRY LEARNING: LOOKING OR DOING?

Anna Shvarts and Linda Zenger

Utrecht University, the Netherlands

Dynamic visualizations have the ambition to show the students the relation between mathematical inscriptions, such as a unit circle and a sine graph. Alternatively, the action-based embodied design approach proposes that the students need to explore and discover this relation based on continuous feedback. In a contrasting multiple-case study, we analyze the opportunities and limitations of those two approaches. We show that higher-performing students can learn well with both design types. Yet, for the lower-performing students, an opportunity to actively discover the relation between the unit circle and sine graph is critical for incorporating this relation into their understanding of the sine graph construction.

INTRODUCTION

Imagine, a student is about to learn how to build a sine graph from the unit circle. A variety of dynamic visualizations have been designed to show the students the relations between a point on the unit circle and a correspondent point on a sine graph (e.g., DeJarnette, 2018). In these technological solutions, as in many other dynamic solutions for mathematics learning, the connection between two representations is already embedded into technology (e.g., Yerushalmy, 1991). Even if a student is free to manipulate one of the visual representations, the other one automatically adjusts to the student's manipulation (Rolfes et al., 2020).

In these educational activities, the students need to observe mathematical relations embedded in a digital environment. Visualizations show something to students and designers expect students to be able to distinguish target aspects of the visualizations and discern their relations (e.g., an arc length and the distance on the x-axis and their correspondence). However, a cultural-historical claim that perception develops (Radford, 2010; Vygotsky, 1997) and a radical embodied claim that perception serves action (Abrahamson & Sánchez-García, 2016; Maturana & Varela, 1992) point out that students need to acquire the ability to noticing the target relations.

Contrary to dynamic visualizations, in action-based embodied designs (Abrahamson, 2014), mathematical relations are delivered in a form of a motor problem (Bernstein, 1967), thus fostering the development of moving in a new way. Embodied interactive activities provide continuous feedback to students' movements similar to an everyday environment (think of a young skater who learns to skate based on continuous feedback from gravity and ice). Aiming at positive feedback, the students learn to maintain visual scene in the target state and develop new sensory-motor coordinations. The core difference between dynamic visualizations and action-based embodied designs can be explained by answering the question, "Who gets to constrain the student's interaction

2023. 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. 203-210). PME 46.

with the virtual objects: the software or the student?" (Abrahamson & Abdu, 2020). By constraining/maintaining new cultural forms of action, students develop new forms of perception: new perceptual structures emerge for them as helpful in maintaining their performance (Abrahamson et al., 2015)

Theoretically, we expect that dynamic visualizations are helpful for students who already have target perceptual abilities and are able to discern target aspects of the visual environment. Action-based designs provide opportunities to develop such perceptual ability within sensory-motor coordinations that students spontaneously develop while exercising within specially designed environmental constraints without "showing". In this study, we aim to investigate *opportunities and limitations in the use of action-based embodied design in comparison with dynamic visualizations*.

THEORETICAL BACKGROUND

In this paper, we exploit an idea of *a body-artifacts dynamic functional system* (Shvarts et al., 2021); this theoretical approach joins and creatively adapts cultural-historical ideas on higher psychological functions as systemic entities mediated by artifacts (Vygotsky, 1997) with radical embodied ideas on cognitive functions as emergent entities within complex dynamic systems of sensory-motor processes in brain and body (Chemero, 2009). The idea of functional dynamic systems comes from physiology and means spontaneous emergent unity of neuronal and peripheral elements (e.g., muscles) activated to fulfill a particular physiological function, such as breathing, or locomotion (Kazansky, 2015). Unlike physiological functions, higher psychological functions (including mathematical thinking) are mediated by cultural artifacts (Vygotsky, 1997). We understand this mediating functionality of the artifacts as a direct extension of a body-brain functional system (Shvarts et al., 2021). While such direct extension is easy to imagine for a fork or scissors when talking about visual or audial notations, we need to take into account the interconnection with the artifacts through the air. Yet, the specificity of eye movements when using mathematical visuals shows that our body is directly involved in operating with mathematical artifacts, just like a hand is involved in operating a fork.

The fundamental mechanism of a functional system lies in continuously anticipating of the environment as it will appear when the body moves (feedforward) and adjusting actions based on feedback that may match or not match anticipation (Bernstein, 1967). Logically, attaching an artifact to the body transforms the horizon of anticipation and moves the border between an agent and an environment to the end of artifact (Shvarts et al., 2021). This way, we anticipate how a fork would behave when moving it. Let us now apply those ideas to trigonometry learning.

Trigonometry knowledge requires interconnecting diverse inscriptions (artifacts) in one system so that they can be flexibly operated: triangle, unit circle, and graphs are in the list of visual inscriptions (Presmeg, 2008); radians and degrees presents two types of notations that incorporate different measurement actions; additionally, algebraic and verbal notations for trigonometric functions need to be acquired.

From the body-artifacts functional system perspective, we expect that the differences between dynamic visualization and an action-based embodied design approach are the following: When students manipulate a dynamic visualization, it extends the students' hand and the student might learn to anticipate its movement as a whole. Technological environments interconnect inscriptions, and students operate with their readily interconnected system. Think of pedalling on a bike: moving legs, we only anticipate the movements of the wheels but are ignorant of interconnections between the pedals, gears, and chain. In the case of embodied action-based design, the students need to build interconnection of different inscriptions themselves. The functionality of interconnection develops within their body.

In this paper, we specifically focus on the sine function on a unit circle, the construction of the sine graph in connection with the unit circle, and finding algebraically expressed sine values on those visualizations As we aim to compare two design genres using the body-artifacts system's idea, our research question is:

What are the differences and similarities in the appropriation of trigonometric inscriptions in the body-artifacts functional systems of the students who learned with action-based embodied designs or dynamic visualizations?

METHODOLOGY

We conducted a contrasting multiple-case study (Miles & Huberman, 1994) comparing the gained knowledge and reasoning of the students who passed through action-based embodied design (ED) and interactive dynamic visualization (DV).

Learning materials

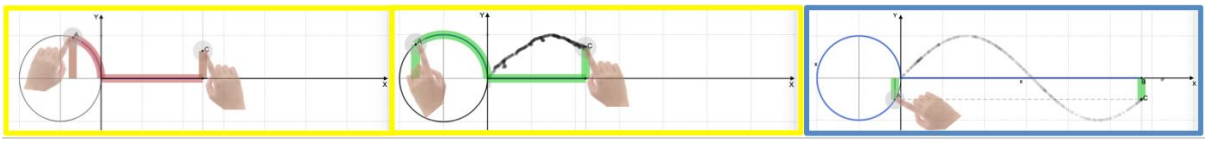
In both conditions, there were four parts: (1) sine value on a unit circle; (2) relation between an arc of a unit circle and the *x*-coordinate of a sine graph; (3) relation between a sine value on a unit circle and *y*-coordinate on a sine graph; (4) constructing a sine graph by relating an arc and a sine value on the unit circle to *x*- and *y*-coordinates of the sine graph (see Figure 1). Each part consisted of (a) an interactive *sensory-motor task*, in which the students studied the relations between a point on a unit circle and a point on a sine graph; (b) a *reflection task* on writing down their observations and then choosing the correct explanations from the multiple-choice tasks, and (c) a *quantification or algebraic task* on estimating the argument and value of the sine function on the graph and unit circle. See details in Shvarts& van Helden (2021).

Contrasting conditions: embodied design and interactive dynamic visualization

Students from ED condition studied the target relations in the form of a motor problem as they manipulated two points: a point on the unit circle and a point on the Cartesian coordinates. Continuous feedback from the system informed them when the points were in the corresponding positions by changing the color from red to green. Students from DV condition moved only one point on the unit circle, while the point on the Cartesian coordinates was automatically moved in correspondence with the first point. In part 4 (shown in Figure 1), in both conditions a line was drawn. In ED condition it

would appear only in case of correct hand movements; in DV condition it would appear automatically when moving around the circle. The reflection and quantification and algebraic tasks were the same in both conditions.

Figure 1. Embodied design, (a) and (b); Dynamic visualization (c)



Participants and research procedure

16 students (14-15 years old) had no pre-knowledge in trigonometry beyond trigonometric relations in a right triangle. The procedure consisted of a pre-test, a learning stage, a post-test, and an interview. In the pre-test, we asked students to estimate the sine value on a unit circle and draw a sine graph. In the post-test, those tasks were repeated (with different values); additionally, we asked students to estimate the sine value on a sine graph. In the interview, explained how they solved the post-test tasks. When needed, we scaffolded them in solving post-test tasks, thus investigating their reasoning. All stages were done online.

Analysis

Firstly, we coded the success in passing tests and performance in each type of task during learning (see Table 1). According to the pre-test, the students in both groups had very limited prior knowledge; based on the post-test results, we distinguished four higher-performing students and four lower-performing students in each condition. Further, we compared the students who showed similar success according to the posttest across different conditions. We analyzed their performance in learning in relation to the learning outcome and multimodal reasoning in the interview. We focused on how they interconnect the sine graph, unit circle, and sine values and include them in their bodily functional systems.

Results

The **higher-performing students** in both conditions could successfully draw the sine graph, explain the construction of its period and amplitude almost without any guidance, and use the sine graph and unit circle for estimating sine values. Interestingly, in both conditions, higher-performing students used gestures while explaining their understanding of the sine graph construction. However, based on their different learning experiences, students used gestures differently. DV students gestured with one hand while verbally explaining the relation between UC and graph. Gestures along the UC were followed by gestures along the graph (i.e., Pier, Cindy) or vice versa (i.e., Pier, Pjotr, Steve). ED students gestured with two hands; they explained the relation between UC and graph by words and also gestured the movement of the corresponding points simultaneously (Lucy, Eva) or subsequently (Erika, Lukas).

Table 1: Performance in tests and learning tasks. Black: full score; white: low score.

	Pre-test		Task Series per three types of the tasks																Post-test	
	Sine	Sine graph	Sensory-Motor task				Reflection task: written & multiple-choice								Algebraic task/ Quantification				Sine	Sine graph
ED:			*1*	*2*	*3*	*4*	*1*	*2*	*3*	*4*	*1*	*2*	*3*	*4*	*1*	*2*	*3*	*4*		
Lukas	0.3	1	3.5	3.5	3	3.5	3	0	3	2	3	2	3	1	1	1	0.8	1	1	3
Eva	0	0	3.5	3.5	3	3.5	1.5	1.5	3	1	1	2	3	3	1	0.7	0.3	0.3	0.8	3.5
Erika	0	0	3.5	3.5	3	3.5	3	2	3	3	1	2	3	0	1	1	1	1	0.9	3.5
Lucy	0	0	3.5	3	3	3.5	2	3	3	3	2	3	3	2	1	1	0.8	0.3	0.8	4
Janice	0	0	3.5	3.5	3	3.5	3	3	3	3	0	0	3	3	0.8	0	0.2	0	0.2	3.5
Jessy	0	0	3	3	2	2.5	2	1	3	1	2	3	3	0	0	1	0.5	0.2	0	2
Said	0	0	3	3.5	4	3.5	3	3	3	2	0	3	2	0	1	1	0.5	0.7	0	3
Jade	0	0	3	3	4	3.5	3	1	3	0	1	0	0	0	0.5	1	0.8	-	-	-
DV:																				
Steve	0	0	-	-	-	-	3	3	3	2	3	3	3	0	1	1	0.7	0.3	1	4
Cindy	0	0	-	-	-	-	3	3	3	3	0	2	2	2	0	1	1	1	0.8	3
Pjotr	0	0	-	-	-	-	2	3	3	3	1	3	2	3	1	1	1	1	1	3.5
Pier	0	0	-	-	-	-	3	3	3	1	1	3	3	2	1	1	1	1	0.7	4
Sem	0	1	-	-	-	-	2	1.5	1	3	0	2	0	3	1	1	1	1	0.6	1
Floor	0	0	-	-	-	-	2	1	1	3	0	0	1	0	1	1	1	0.5	0.3	1
Lola	0	0	-	-	-	-	3	0	2	1	0	0	0	0	0.8	0	0.5	0	0.2	1
Fema	0	0	-	-	-	-	0	1.5	3	0	0	3	1	0	0	1	0.5	0.5	0	1
Max:	1	4	4	4	4	4	3	3	3	3	3	3	3	3	1	1	1	1	1	4

For example, Pier (DV) supported his verbal explanation with gestures: *Uh... well I just knew that, if uh uh... at around half of pi [gesturing around top half UC], when the x-axis [gesturing x-axis], it would reach the top [gesturing], so that would be 1.*

Figure 2. Pier (DV) at first gestured out the unit circle and then the sine graph.

Lucy (ED) relied on the gestures to explain her drawing: *Uh I drew it uh... regarding the points A and C that went up the circle like this [gesturing both points simultaneously for half circle/period] ...*

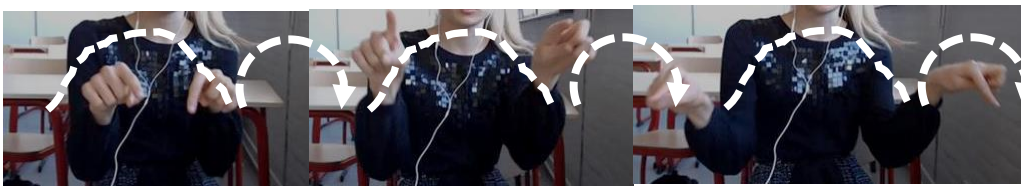
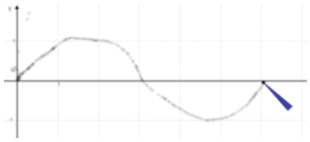
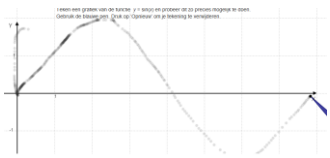
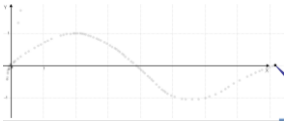
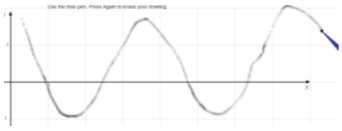
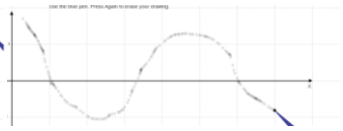
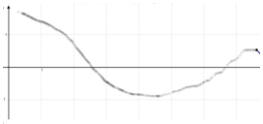
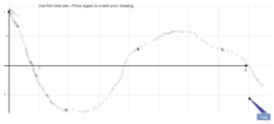


Figure 3. Lucy (ED) gestured the unit circle and sine graph simultaneously.

The post-test results of the **lower-performing students** are quite contrasting for DV and ED students (see Table 1): while ED students seem to perform better in drawing a sine graph, DV students tend to be better in the tasks of estimating sine value based on the unit circle and sine graph. Table 2 presents sine graphs drawn in the post-test by the lower-performing students in each group. Apart from Jade (who never reached the post-test), all ED students grasped that the graph starts from the origin and that it has equal waves up and down; Janice and Said grasped the amplitude correctly, and Janice also grasped the period. All DV students drew some kind of general waves, without specifically caring about amplitude and period.

Table 2. Sine graph in a posttest drawn by students from ED and DV groups

Janise (ED)	Jessy (ED)	Said (ED)	Jade (ED)
			
Simon (DV)	**Floor (DV)**	**Lola (DV)**	**Fema (DV)**
			

In their explanations of graph construction, ED students referred to the unit circle. For example, although not fully correct, Said said: "*I knew the circle passed about half arc at two*". The DV students lost the connection with the unit circle, e.g., Lola and Floor directly said that they *tried to copy* the line that appeared in dynamic visualization without intruding on how exactly it was constructed.

Analysis of the learning process might shed light on these differences in post-test and interviews. Janice and Said were the most successful (among lower-performing students, see Table 1) in their written reflections on the relation between the point in the unit circle and the sine graph. These and other ED students refer to their actions and what they should do. For example, Said (ED) wrote: "I keep the points at the same height, next to each other"; "I care that two points would be at the same distance in comparison to each other from the cross point of the circle and the line". DV students reported on the relation of the points detached from their experience or honestly stated that they could not grasp the relation (never appeared in ED). Lola (DV): "Point S is on the same height as point A"; Floor "no, I don't really have an explanation"; "point C equals the height of A, and point B equals the distance that A has traveled".

DISCUSSION AND CONCLUSIONS

The differences between design types are firstly vivid for low-performing students. ED students had an opportunity to discover the connections between the visualizations in their sensory-motor actions. As we expected, those connections were part of their personal experience (as interviews evidenced) and formed their functional bodily

systems. In DV condition, the connection between inscriptions was initially embedded into the environment, and, as students' reflections evidenced, stayed outside of their body-artifacts systems: their post-test sketches and explanations did not expose any interconnection with unit circles. Sometimes, students reported difficulties in perceiving interconnection (as theories of enculturated perception would expect; Radford, 2010); in other cases, they did not follow their own descriptions while drawing. As interviews evidence, students tended to mimic the system, thus grasping an artifact (the sine graph) as ready-made, without understanding its construction.

At the same time, the DV version appeared to be beneficial for quantifying sine values. ED students were drawn into exploring the interconnection between different visual relations and could not make the next step in distinguishing which distance needs to be estimated as an argument or value of the sine function. Contrary, DV students could step away and operate with UC and the sine graph as ready-made entities and build on them towards further algebraization. Overall, the choice between those two designs particularly matters for the students who might struggle with mathematics and might be determined by an educational aim: dynamic visualizations might be more helpful in teaching to use ready-made mathematical artifacts; embodied designs might help in questioning and re-discovering the construction of mathematics.

At the same time, some students could understand the sine graph construction and learn to estimate sine values on the unit circle and sine graph independently from the design type. Theoretically, we could say that visual inscriptions and their interconnection were appropriated into the body-artifacts functional systems of higher-performing students and could take part in their further mathematical reasoning. Remarkably, the way of learning (ED or DV) fostered different involvement of the body in reasoning: ED students would directly interconnect two artifacts by their bodies (two hands gestures), while DV students would switch between operating with two artifacts through speech.

For ED students, the interconnection of the unit circle and sine graph was established as new motor coordination of two hands (Abrahamson & Sánchez-García, 2016). How, theoretically, could this interconnection have been uncovered by the functional bodily systems of DV students who did not manipulate both inscriptions? We assume that DV students could establish this interconnection coordination between a hand and an eye: they needed to find a way to anticipate the dynamics on the sine graph while manipulating the point on the unit circle with their hand.

Our study certainly has a limitation of a small sample and may only be considered as hypothesis generating. Yet, we think that such a comparison makes a step in specifying the potential of educational technology beyond interactive dynamic visualizations (Rolfes et al., 2020). Embodied action-based design (Abrahamson, 2014) might help in overcoming the limitations of the multiple representations software that require essential teaching effort to support conceptual understanding (Yerushalmy, 1991).

Acknowledgments: We thank Rosa Alberto for the collaboration on the design sequence. The research was supported by the NWO-652.001.004 grant.

REFERENCES

Abrahamson, D. (2014). Building educational activities for understanding: An elaboration on the embodied-design framework and its epistemic grounds. *International Journal of Child-Computer Interaction*, *2*(1), 1–16.

Abrahamson, D., & Abdu, R. (2020). Towards an ecological-dynamics design framework for embodied-interaction conceptual learning: The case of dynamic mathematics environments. *Educational Technology Research and Development*, 1–35.

Abrahamson, D., & Sánchez-García, R. (2016). Learning is moving in new ways: The ecological dynamics of mathematics education. *Jour. of the Learn. Sci.*, *25*(2), 203–239.

Abrahamson, D., Shayan, S., Bakker, A., & van der Schaaf, M. (2015). Eye-tracking Piaget: Capturing the emergence of attentional anchors in the coordination of proportional motor action. *Human Development*, *58*(4–5), 218–244.

Bernstein, N. A. (1967). *The co-ordination and regulation of movements*. Pergamon press.

Chemero, A. (2009). *Radical embodied cognitive science*. MIT Press.

DeJarnette, A. F. (2018). Students' conceptions of sine and cosine functions when representing periodic motion in a visual programming environment. *Journal for Research in Mathematics Education*, *49*(4), 390–423.

Kazansky, A. B. (2015). Agential anticipation in the central nervous system. In M. Nadin (Ed.), *Cognitive Systems Monographs* (Vol. 25, pp. 101–112). Springer, Cham.

Maturana, H. R., & Varela, F. J. (1992). *The tree of knowledge: The biological roots of human understanding*. Shambhala Publications.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative Data Analysis: An expanded soursebook*. SAGE Publications.

Presmeg, N. (2008). Trigonometric connections through a semiotic lens. In L. Radford, G. Schubring, & F. Seeger (Eds.), *Semiotics in Mathematics Education: Epistemology, History, Classroom, and Culture* (pp. 103–120). Sense Publishers.

Radford, L. (2010). The eye as a theoretician: Seeing structures in generalizing activities. *For the Learning of Mathematics*, *30*(2), 2–7.

Rolfes, T., Roth, J., & Schnotz, W. (2020). Learning the concept of function with dynamic visualizations. *Frontiers in Psychology*, *11*.

Shvarts, A., Alberto, R., Bakker, A., Doorman, M., & Drijvers, P. (2021). Embodied instrumentation in learning mathematics as the genesis of a body-artifact functional system. *Educational Studies in Mathematics*, *107*(3), 447–469.

Shvarts, A., & van Helden, G. (2021). Embodied learning at a distance: From sensory-motor experience to constructing and understanding a sine graph. *Math. Think. and Lear.*, 1–29.

Vygotsky, L. S. (1997). The history of the development of higher mental functions. In R. W. Reiber (Ed.), *The Collected Works of L. S. Vygotsky* (Vol. 4). Springer US.

Yerushalmy, M. (1991). Student perceptions of aspects of algebraic function using multiple representation software. *Journal of Computer Assisted Learning*, *7*(1), 42–57.

STRATEGY USE IN NUMBER LINE TASKS OF STUDENTS WITH AND WITHOUT MATHEMATICAL DIFFICULTIES: A STUDY USING EYE TRACKING AND AI

Anna Lisa Simon[1], Parviz Asghari[1], Achim J. Lilienthal[2], & Maike Schindler[1]

[1]University of Cologne, [2]Technical University of Munich and Örebro University

The number line (NL) is an important tool in mathematics education. Students with mathematical difficulties (MD) tend to have difficulties in NL tasks, as indicated by qualitative analyses of eye-tracking data. However, these qualitative analyses are laborious especially for large amounts of data. Our paper uses an innovate approach to facilitate the analysis of student strategies: AI is used to support the human researchers. We use student gaze heatmaps in combination with AI, in particular a clustering algorithm, to identify strategies of 140 fifth-grade students on the NL. Through AI-enhanced analysis, we found, first, a set of student NL strategies different from previous research. Second, we found that students with and without MD—at certain numbers—differed in strategy use, which was not found in this way before.

INTRODUCTION

Central to learning arithmetic in primary school is to develop a number sense, including an ordinal understanding of numbers (Fuson, 1988). To this end, the number line (NL), where numbers are represented in a linear arrangement, is often used (Diezmann & Lowrie, 2007). For being able to use the NL, it is necessary to understand and interpret numbers in relation to each other. It has been shown that students' performance on NL tasks is closely related to their overall mathematical achievement (Schneider et al., 2018), which highlights the significance of the NL for students' number sense.

Mathematical difficulties (MD) are an important topic in mathematics education (research). Students with MD typically show difficulties in learning basic arithmetic at the primary school level that can persist into secondary school (e.g., Moser Opitz et al., 2017). Previous research has shown that students with MD tend to have difficulties in NL tasks, such as lower accuracy in locating numbers and lower flexibility in using strategies (e.g., van't Noordende et al., 2016). Eye tracking (ET), in particular, the analysis of ET videos, has been shown to be for investigating student strategies on the NL (Simon & Schindler, 2022). However, since qualitative analyses of ET videos can be laborious, more efficient analysis methods are needed, especially for large amounts of data. Research has shown that Artificial Intelligence (AI) can help identifying student strategies from ET data, particularly from gaze heatmaps, which are visual representations displaying gaze distribution and intensity (Schindler et al., 2020, 2022).

The aim of this paper is twofold: First, we investigate fifth-grade students' strategy use in locating numbers on a NL. Second, we investigate differences of students with MD as compared to students without MD in their use of NL strategies. We use AI-enhanced analysis of gaze heatmaps, specifically a clustering algorithm (unsupervised machine

2023. 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 46.

learning) to find strategies. This approach provides an independent perspective on the data. For a qualitative analysis the AI suggestions were interpreted by a human expert who can verify meaningful student strategies. For analyzing group differences in strategy use, we analyze how the strategies are represented in the cluster analysis.

THEORETICAL BACKGROUND

Number line

The NL is a fundamental tool in teaching and learning mathematics, particularly at the primary school level. As it represents the basic idea of the number series, it is used to develop number sense and to deepen an ordinal understanding of numbers (Diezmann & Lowrie, 2007). In addition to the ordinal understanding of numbers, the relational interpretation of numbers is an important way to make numbers accessible on the NL. There are different types of NL, e.g., those that have only labeled beginning and endpoints (empty NL) and others that have more hatch marks, e.g., for ones or tenths represented on the NL (marked NL). NL have changeable elements (e.g., different ranges of number represented in NL) so that different interpretation of distances and hatch marks on the NL is required (Teppo & van den Heuvel-Panhuizen, 2014). Thus, students need to interpret numbers on the NL according to the given structuring features. The NL can be used in different ways and needs to be interpreted and filled with mathematical meaning, which can cause difficulties for students, especially if not all numbers on the NL are marked and labeled (Schulz & Wartha, 2021). Studies have found that school-age students' performance with NL is correlated with overall mathematical achievement, indicating that students who struggle with mathematics may have difficulties using the NL as well (Schneider et al., 2018).

Mathematical difficulties

Not least since implementation of the inclusive school system, MD have been an important topic in practice and research. MD are characterized by difficulties in understanding basic arithmetic concepts (e.g., Moser Opitz et al., 2017), e.g., in basic quantity-number competencies such as developing basic ideas of numbers and operations. These difficulties typically become apparent at the beginning of primary school, can persist into secondary school level (e.g., Moser Opitz et al., 2017), and can lead to difficulties even beyond school years, e.g., in job contexts. To support students with MD adequately, it is necessary to identify students' individual strengths and difficulties in mathematics. Students with MD tend to have difficulties, among others, in NL tasks. Previous studies have shown that students with MD tend to be less accurate in locating numbers on the NL and tend to have difficulties in the use of adequate strategies as compared to students without MD (e.g., van't Noordende et al., 2016).

Eye tracking and the use of AI to support analysis of eye-tracking data

ET, the recording of eye movements (Holmqvist et al., 2011), has been used in several studies in mathematics education research (Strohmaier et al., 2020). Analyzing ET data has been shown to provide information about student strategies in different

mathematical tasks, and ET has also been shown to be useful in analyzing students' strategies in NL tasks (e.g., van't Noordende et al., 2016). In a previous ET study using marked NL tasks similar to the present study, ET videos (i.e., videos where gazes are visualized as a semi-transparent dot) were analyzed qualitatively (Simon & Schindler, 2022) and it was found that students use counting or direct strategies. However, qualitative analysis of such gaze patterns is challenging and time-consuming, and requires advanced domain knowledge. To reduce the effort of qualitative analysis, we use AI, specifically a clustering algorithm based on gaze heatmaps as input. Previous research has already shown how differences in strategy use for enumeration tasks between different groups of students can be analyzed using cluster analysis (Schindler et al., 2022). It showed how AI can work together with human researchers as an AI colleague making suggestions about ways to categorize the data based on similarities in appearance, while the human experts then interpret and verify these suggestions.

We ask the following research questions: (1) *What strategies for locating numbers on a marked NL do fifth grade students use?* (2) *Do students with MD differ from students without MD in their use of these strategies?*

THIS STUDY

Participants. The study took place at a German comprehensive school at the beginning of fifth grade. All students participated in a standardized arithmetic test, the HRT (Haffner et al., 2005), to diagnose MD. Students whose performance shows a PR≤10 are classified as having MD (*with MD*). Students whose performance shows a PR>25 are classified as not having MD (*without MD*). Students with a PR of 11-24 are considered "at risk" for MD. Based on the design of the NL and the distance between hatch marks, we chose to exclude students with ET data accuracy ≥2° from further analysis, so *N*=140 students (mean age: 10.6 years, SD: 0.6 years) were considered.

Tasks, procedure, and eye tracker. We used a marked NL with two different ranges of numbers. For position-to-number-tasks (Fig. 1), where students are shown a position on the NL and asked for the corresponding number, we used the numbers 40, 55, and 70 on NL 0–100, and the numbers 250, 650, and 800 on NL 0–1000. Tasks were presented on a computer screen in the same randomized order for every student.

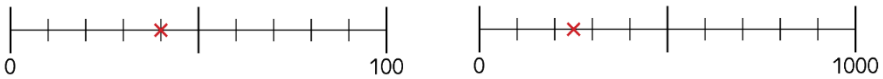


Figure 1. Examples of NL tasks (number 40, left; number 250, right).

Students were tested in a quiet room at school. Each student sat about 60 cm in front of a 24'' full HD monitor. Students received instructions through a headset, but also had the opportunity to ask questions if they did not understand instructions. Before the above-mentioned tasks, the students were shown an instruction video and a trial task to ensure they understood the instructions. Students were asked to type the correct number on a number pad and, in between the tasks, to fixate a star in the upper left corner of the screen. The students received no feedback whether their answers were correct. To record students' gazes, we used the remote eye tracker Tobii Pro X3-120

(infrared, binocular, 120 Hz). The eye tracker was calibrated for each student before the work on the tasks. The average accuracy of the ET data was 0.90° (SD = 0.31°).

Heatmaps and clustering. In general, student gaze patterns may differ in several ways, e.g., in terms of locations, durations, and the order of gazes on the NL. For the analysis of student strategies, we chose heatmaps as input (Holmqvist et al., 2011). Heatmaps are visual representations that display the spatial distribution of gazes for each task. They represent the duration of gazes on certain areas in relation to the total duration of the task but no sequential information. In our study, heatmaps were generated based on raw data provided by the eye tracker and the noise was technically reduced with a method presented by Asghari et al. (2023) to facilitate clustering of heatmaps.

In this paper, we use ET in combination with AI to investigate students' strategies in NL tasks (similar to Schindler et al., 2020, 2022). In particular, we used unsupervised machine learning which is a set of methods that tries to "find 'interesting patterns' in the data" (Murphy, 2012, p. 2). In particular, we used a clustering algorithm, which separates the data (here: gaze heatmaps) into a number of meaningful clusters. As clustering algorithm, we used Self-Organizing Maps (SOMs, Kohonen, 2001) (for a detailed description, see Schindler et al., 2020, 2022). Briefly stated: In the clustering process, each heatmap is assigned together with similar heatmaps to a cluster. Since the number of clusters is predefined, clusters can remain effectively empty. Given that the heatmaps in each cluster are similar in appearance, we assume that the different clusters represent specific strategies. We computed average heatmaps as prototypical examples of strategies for each cluster. As similarity metric, we use the cosine distance, which measures the similarity between the direction of two vectors, regardless of their magnitude. This metric is used, e.g., in image processing to measure the similarity between two documents, or two images.

Qualitative analysis of clusters. Using SOMs, the gaze heatmaps were separated into a maximum of twelve clusters for each of the six tasks ("AI-colleague"). For each of the twelve clusters per task, we then interpreted the average heatmaps ("human expert") (Fig. 2). We assigned strategies to the average heatmaps based on a category system inductively developed in a previous study (Simon & Schindler, 2022) with six different strategies for locating numbers on a marked NL. To rule out that students just guessed the answers without considering the given information seriously, only heatmaps of correctly solved tasks were considered for analyses. This resulted in slightly different numbers of considered heatmaps for each task. For clustering, we used heatmaps of all fifth-graders to have the largest possible data set for identifying NL strategies (RQ 1). The average heatmaps of clusters with at least five heatmaps were then assigned a strategy. Clusters with fewer heatmaps were not considered, since a meaningful interpretation was not possible due to the possibly low similarity of individual heatmaps within these clusters (clusters not shown in Fig. 2). For statistical analyses of possible differences in strategy use between students with and without MD (RQ 2), students "at risk" for MD were not considered, since they cannot be assigned either of the two groups.

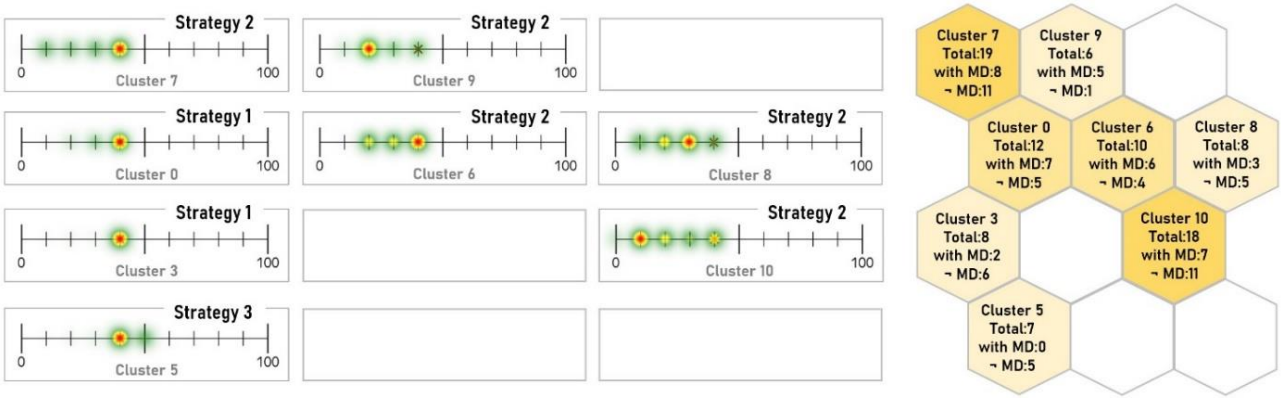


Figure 2. Clusters for number 40 and strategy assignment to average heatmaps (¬MD represents students without MD).

Statistical analysis. To investigate differences in strategy use between students with and without MD, we conducted chi-squared tests using SPSS 29. For chi-squared tests that showed significant group differences, we calculated cell tests to examine group differences in more detail. Effect sizes were calculated using *Cramérs V*.

RESULTS

To answer research question 1 (*What strategies for locating numbers on a marked NL do fifth grade students use)*, we investigated student strategies based on the clustering of heatmaps. We found four strategies that reflected the orientation of the students on the NL (examples of average heatmaps for strategies shown in Fig. 3).

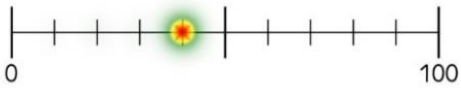
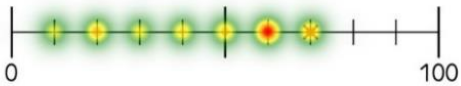
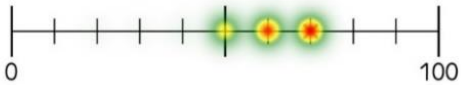
Strategy number	Strategy name	Example average heatmap	Corresponding number
Strategy 1	Direct orientation		40
Strategy 2	Orientation from the beginning		70
Strategy 3	Orientation from the midpoint		70
Strategy 4	Orientation from the endpoint		800

Figure 3. Examples of the strategies visualized by average heatmaps.

We found that the number of strategies used for each task varied in the following way: (I) For 250 on the NL, we found strategies 1 and 2. (II) For 40, 55, and 650 on the NL, we observed strategies 1, 2, and 3. (III) For 70 and 800 on the NL, we found strategies 1, 2, 3, and 4 that were used by the students. Based to these results, tasks with the same number of strategies used by the students were combined for the statistical analyses.

To answer research question 2 (*Do students with MD differ from students without MD in their use of these strategies?*), chi-squared tests were conducted. There was no difference in strategy use between students with and without MD for (I) number 250 (χ^2 (1) = .03, p = .864), and also for (II) numbers 40, 55, and 650 (χ^2 (2) = .49, p = .785). Chi-squared test for (III) numbers 70, and 800 revealed significant differences in the use of strategies between students with and without MD, with medium effect size: χ^2 (3) = 14.70, p = .004, V = .28. In detail, cell tests revealed that students with MD used orientation from the beginning (i.e., strategy 2) significantly more often (small effect size) than students without MD (χ^2 (1) = 5.89, p = .015, V = .18), and that students without MD used orientation from the endpoint (i.e., strategy 4) significantly more often (small effect size) than students with MD (χ^2 (1) = 11.20, p = .003, V = .24) (Bonferroni-Holm adjusted p-values).

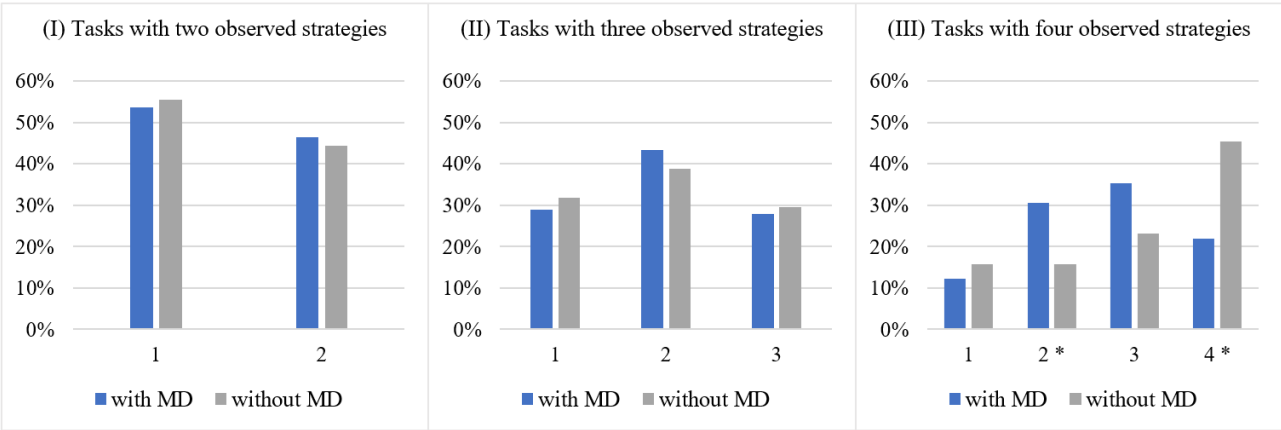


Figure 4. Students' strategy use for conditions (I), (II), and (III) (significant differences are marked with *).

DISCUSSION

The aim of this paper was to investigate (1) fifth-grade students' strategy use in locating numbers on a marked NL and (2) differences of students with MD as compared to students without MD in their use of these strategies. To pursue this aim, we analyzed gaze heatmaps with the help of AI, in particular a clustering algorithm. We investigated students' strategy use based on average heatmaps representing different NL strategies.

(1) We found a set of NL strategies different from previous findings: In previous studies, strategies on a marked NL were mainly classified based on whether the strategies involved counting or rather direct attention to certain points (Simon & Schindler, 2022). Through our AI-enhanced analysis the strategies were clustered rather by the areas of the NL that were attended to most, which relates to the reference points the students used for orientation. Although previous research has looked into reference points in students' strategies also, in this study the use of reference points was emphasized much more through the AI-enhanced heatmap analysis. This might be due to the fact that heatmaps represent the gaze distribution and intensity, but no temporal information about students' gazes, i.e., about their order. Looking at the set of four strategies found through the help of AI, it is interesting that they were not only

meaningful from a mathematics education perspective, but also categorized along cases of reference points as known from previous research: direct orientation, or orientation from the beginning, midpoint, or endpoint (Simon & Schindler, 2022). Looking into the use of strategies found by help of AI, it is also meaningful that for numbers at the beginning of the NL (here: 250), only two strategies were found (i.e., direct orientation and orientation from the beginning), while for numbers at the end of the NL (e.g., 800), also the midpoint and endpoint of the NL were used for orientation. This shows that the AI-enhanced research design, where AI provided clusters that were interpreted by a human expert, provided insightful results that were meaningful from a mathematics education perspective and contributed to the state of research.

(2) We found significant differences in strategy use between students with and without MD—even for the small number of different NL strategies. These significant differences occurred only for specific numbers on the NL: for numbers near the endpoint of the NL, i.e., 70 on NL 0-100, and 800 on NL 0-1000. Students with MD, in order to locate numbers near the endpoint of the NL, were more likely to be oriented towards the beginning of the NL, and less likely to be oriented towards the endpoint of NL as compared to students without MD. This is in line with previous findings on the difference in strategy use in empty NL tasks between students with MD and without MD (e.g., van't Noordende et al., 2016). New to our findings on the orientation of students on a marked NL is that especially the orientation towards the endpoint of the NL of students without MD appears to make the difference to students with MD, who less likely use the end of the NL for orientation purposes.

In addition to these empirical findings, our results indicate that the use of AI for providing an independent view on ET data by making category suggestions is a promising tool for supporting human researchers in mathematics education in the analysis of gaze heatmaps.

ACKNOWLEDGEMENT

This project has received funding by the Federal Ministry of Education and Research as a part of the program KI-ALF (01NV2123). The responsibility for the content of this publication remains with the authors.

REFERENCES

Asghari, P., Schindler, M., Simon, A. L., & Lilienthal, A. J. (2023, ETRA, under review). Pre-processing eye-tracking data with domain-knowledge-aware dimensionality reduction through automatic offset calibration and gaze remapping to improve the performance of downstream data analysis.

Diezmann, C., & Lowrie, T. (2007). The development of primary students' knowledge of the structured number line. In *Proceedings of PME31* (Vol. 2, pp. 201-208). PME.

Fuson, K.C. (1988). Children's counting and concepts of number. Springer.

Haffner, J., Baro, K., Parzer, P., & Resch, F. (2005). *Heidelberger Rechentest*. Hogrefe.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., and Van de Weijer, J. (2011). Eye tracking: A comprehensive guide to methods and measures. Oxford UP.

Kohonen, T. (2001). *Self-Organizing Maps*. Springer.

Moser Opitz, E., Freesemann, 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

Murphy, K. P. (2012). *Machine Learning: A Probabilistic Perspective*. The MIT Press.

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., Schaffernicht, E., & Lilienthal, A. J. (2020). Identifying strategies through eye tracking and unsupervised learning: The case of quantity recognition. In *Proceedings of PME44* (pp. 518-527). PME

Schneider, M., Merz, S., Stricker, J., De Smedt, B., Torbeyns, J. Verschaffel, L., & Luwel, K. (2018). Associations of number line estimation with mathematical competence: A meta-analysis. *Child Development, 89*, 1467-1484.

Schulz, A., & Wartha, S. (2021). *Zahlen und Operationen am Übergang Primar-/ Sekundarstufe. Grundvorstellungen aufbauen, festigen, vernetzen*. Springer.

Simon, A. L., & Schindler, M. (2022). Strategy use in number line tasks of students with mathematical difficulties: An eye-tracking study. In *Proceedings of PME45* (Vol. 4, pp. 3-10). PME.

Strohmaier, A. R., MacKay, K. J., Obersteiner, A., and Reiss, K. M. (2020). Eye-tracking methodology in mathematics education research: A systematic literature review. *Educational Studies in Mathematics, 104*, 147–200.

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.

A NOVICE TEACHER'S IDENTITIES – FROM LOSING HER BALANCE TO REGAINING HER CONFIDENCE

Charlotte Krog Skott[a], Jeppe Skott[b, c]

[a]University College Absalon, Denmark;

[b]Linnaeus University, Sweden; [c]University of Agder, Norway

Novice mathematics teachers often find the initial years of classroom teaching demanding. In this paper, we follow a novice teacher, Laura, the first 18 months of her teaching career and investigate the development of her professional identities from a participatory perspective. Laura's case is interesting as her identities change dramatically, due to her participation in an induction programme and to a change of school. Our analysis shows that she loses her professional balance and almost leaves the profession during the first year, and that she regains her professional confidence and even coaches her colleagues on reform-oriented teaching in the second year. We argue that our participatory approach helps explain Laura's identity trajectory.

INTRODUCTION

Professional identity is a significant field in research on and with teachers (Darragh, 2016; Lutovac & Kaasila, 2018). In comparison with for instance research on teachers' knowledge and beliefs, identity studies tend to adopt a more participatory stance that moves beyond individual cognition as the main focus of attention (Skott, 2022). Like other participatory fields, it "seek[s] to place thinking agents in their larger social, physical, cultural, and historical contexts" (Russ et al., 2016, p. 403). It follows, that identities are seen as multiple, fluctuating and contextually dependent.

Our study is in line with this approach. We followed a Danish lower secondary teacher, Laura, for the first 18 months of her career. She got her first job at a municipal school, Southbank, at which she was challenged by resistance from students and parents, by lack of leadership support, and by limited collegial collaboration. After six months, Laura became involved in an induction programme in which she worked with two other novices, their school mentors, and three teacher educators (including the first author). At the end of the year, she got a job at nearby a private school, Blackheath. Our aim is to describe and explain Laura's identity trajectory over these 18 months.

Defining Laura's identities as her professional experiences of being, belonging and becoming and using a framework called Patterns of Participation (PoP), we address the questions of (1) how Laura's identities change over the first 1½ years of her career and (2) what seems to promote these changes.

LITERATURE: IDENTITY – EMOTION AND RETENTION/ATTRITION

Much of the scholarship on professional identities focus on novice teachers' challenges with their new profession (e.g. Jong, 2016; Skott, 2019). Pillen et al. (2013) point to how conflicts between professional requirements and "what they personally desire or

2023. 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. 219-226). PME 46.

experience as good" may become emotional challenges to their professional learning and reciprocally related to their identity development (p. 660-661). Based on qualitative interviews, Pillen et al. found three main themes in the tensions the participants experienced, namely those between (1) being a student and a teacher, (2) desired and actual support to students, and (3) teaching-learning processes favoured by the participants themselves and by their mentors. Participants learnt from the tensions, but often reacted with "helplessness, frustration, or anger" (p. 674).

Another but related theme in the literature links identity to teachers' decision to stay in the profession or not (e.g. Cochran-Smith et al., 2012; Hong, 2010). This question of teacher retention and attrition "*refers to the need to prevent good teachers from leaving the job for the wrong reasons*" (Kelchtermans, 2017, p. 965, emphasis in original). Arguing that teaching depends on contextual conditions, Cochran-Smith et al. (2012) describe possible connections between teacher quality and teachers' propensity to stay or leave. As aspects of teacher quality, they include the character of the classroom environment, "the scope, sequence, and substance of what is taught" (p. 854), and "the richness and cognitive complexity" of students' learning opportunities (p. 856). Based on a multiple case study, they found five constellations of practice and career decisions, for instance one that combines high teacher quality and a decision to stay and another in which weak teachers move to a new school.

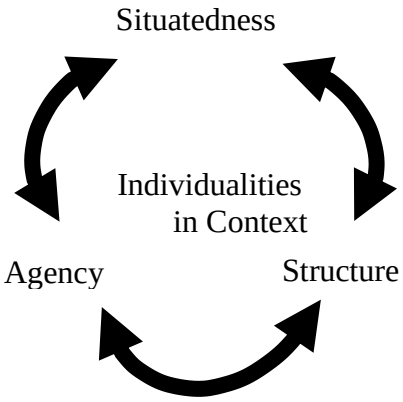
As we shall see, emotional tensions and questions of quality instruction are prominent aspects of Laura's professional experiences at the beginning of her career. We interpret these experiences with reference to social and cultural perspectives on the profession not unlike the ones mentioned by Cochran-Smith et al. However, we shift the focus from *teacher* quality to *teaching* quality by means of the PoP framework.

THE POP FRAMEWORK

Like other participatory fields, most identity research takes the individual-in-social-practice as its unit of analysis. There are multiple interpretations of what this means in studies of identity. However, Skott (2022) has suggested that there is a core to the identity construct as used in field, an *identity triad* consisting of structure, situatedness and agency, which may be used to distinguish between different approaches to *Individualities in Context* (Figure1).

In the identity triad, structure concerns issues that are external to the current situation (e.g. public discourses on education; administrative/political decisions). Approaches that foreground situatedness emphasize what and how professional identities emerge as teachers participate in local practices at a school or in a professional development programme. Identity studies that emphasise agency, acknowledge that structures and local situations orient teachers' actions, but "do so in open-ended ways, leaving space for professional decision-making and agency" (Skott, 2019, p. 470).

PoP is located near the link between agency and situatedness. It focuses on the emergence of identities in the locally social (e.g. Skott, 2019) by drawing on symbolic interactionism (SI) (e.g. Blumer, 1969) and social practice theory (e.g. Holland et al., 1998; Lave, 2019; Wenger, 1998). Relevant constructs from SI include those of interaction and self. As people interact, they see themselves from other's perspectives, including the perspectives of generalised others such as communities and discourses. Key constructs from social practice theory include those of practice and figured worlds (FWs). An FW is a social and cultural "realm of interpretation, in which particular characters and actors are recognized, significance is assigned to certain acts, and particular outcomes are valued over others" (Holland et al., p. 52). PoP is not oblivious to the significance of structural issues and argues that they may function as FWs that play the role of generalized others in interaction. In the process, teachers' professional experiences of being, becoming and belonging may change. In identity research, PoP is to shed light on how and why this is so.



igure 4. The Identity Triad, from Skott (2022)

METHODOLOGICAL APPROACH

We selected Laura for the study because of marked changes in her professional identities during and after her teacher education programme. She enrolled in the programme almost by accident and chose mathematics because she did not want to do Danish or in English, the only alternatives. However, she developed an interest in mathematics in the course of the programme and gained confidence in teaching it.

Danish teacher education for primary and lower secondary school is a 4-year Bachelor's programme. Prospective teachers specialize in three subjects, one of them being Danish, English or mathematics. The mathematical specialization includes the subject itself and the related educational issues, and it models teaching-learning processes (e.g. inquiry) that teachers are expected to initiate upon graduation.

In Denmark, no systematic support is provided to new teachers. Therefore, the first author and her colleagues designed a one-year induction programme for mathematics teachers inspired by lesson study (cf. C. K. Skott et al., 2021). In 2019, Laura participated in the programme together with two other females, all of whom graduated recently from prestigious Danish colleges. They participated in:

1) *A pre-interview* (1 hour) (referred to as [1]) that focused on their challenges as novice teachers; they emphasised inquiry-based teaching as a significant challenge;

2) *Two individual lesson studies*, in which each novice teacher planned and conducted inquiry-based lessons with her mentor and the teacher educators;

3) *One joint lesson study* in which the three teachers, the mentors and the teacher educators jointly planned a lesson on *What natural numbers can be written as a sum of consecutive natural numbers?* The teachers took turns teaching the lesson, observed by her mentor and the educators. Finally, all participants met to revise the plan ([2]).

Our data material consists of video- and audio-recordings of the above activities. In addition, we conducted two individual interviews with Laura ([3] and [4] respectively), one in 2019 after the pre-interview and one in 2022. These interviews focused on her experiences from the teacher education programme and from working at Southbank and Blackheath.

We conducted the analysis in two steps. In the first step, inspired by grounded theory (Charmaz, 2014), we empirically inferred the figured worlds and practices that played prominent roles for Laura's experiences. We constructed initial codes, for instance 'Using tools to regulate student behaviour'. Comparing our initial codes, we constructed focused codes, such as 'Managing classrooms by regulating behaviour', which we assembled into different figured worlds, e.g. one of 'Teaching mathematics as regulating student behaviour', which was prominent at Southbank. Similarly, we constructed a world of reform-oriented teaching, which encompassed the practices, characters and valued outcomes related to inquiry-based teaching that Laura encountered at college.

In the second step, we used the figured worlds and practices resulting from the first step as an analytical lens to explore Laura's initial experiences and identity formation. We did so for two separate phases, her first year at Southbank, which we call *A lonely struggle – losing professional balance*, and her the second year at Blackheath, called *Regaining confidence – merging structure and inquiries*.

RESULTS AND ANALYSIS

A lonely struggle – losing professional balance

Although educated as a lower secondary teacher, Laura taught mathematics in grade 3 at Southbank. Often, Danish students have the same mathematics teacher in grades 1 to 3, but Laura was the fourth mathematics teacher of her class. She experienced the classroom atmosphere as harsh and referred to the students' attitude to the teachers as a matter of "seeing who we [the students] can kick out the fastest" [3]. Initially, Laura drew on reform-oriented ideas from her pre-service education, but she experienced resistance from the students: "I have many good ideas on paper, but in the classroom … they turn out to be really bad" [1]. Gradually, she yielded and entered the classroom with a plan A, consisting of investigative tasks, and a plan B, based on textbook exercises. However, "the more playful approach to mathematics [plan A] disappeared rather quickly" [3]. Increasingly, Laura distanced herself from the reform-world and criticized it for not providing sufficient structure for the tasks and students' activities.

Laura also experienced her interactions with the parents as challenging. They refused to recognize problems with their children and accused her of being too inexperienced.

At one point, she says, "they even reported me to the police" [4]. However, Laura's older colleagues had similar problems with the class and together they asked the leadership for help. For a period, they even refused to teach the class without further support, but the leadership ignored them and did nothing to address the problems.

Laura turned to her colleagues. They tried to help, but did so by suggesting how to structure the classroom by applying regulating tools, which kept the students calmer, but positioned them as more passive learners. In other terms, the suggestions helped Laura solve disciplinary problems and make the students stay on task, but did so by drawing on a world of regulating student behaviour that dominated Southbank. However, she lost her passion for teaching and experienced not being good at the job or becoming better. She began to blame herself and asked "is it because I am not able to do it at all [teach mathematics]?" [4]. She expected her first year as a new teacher to be hard, but realised that for her "it has been quite a bit harder than the norm … there were months, when I cried […] every day" [4]. The result of it all was, that Laura experienced not being valued as a teacher at Southbank and not belonging to a community of colleagues, as the relations were only professional.

In relation to her professional belonging, Laura's participation in the induction programme became important. She found that she was not the only one, who struggled with disciplinary problems, who had given up on inquiry-based teaching and who felt left to fend for herself. She experienced belonging to this group of novice teachers.

At the end of the year, Laura turned down job interviews, because she was afraid to experience yet another set of failures under new circumstances, which would force her to accept that *she* was the failure. However, she, somewhat reluctantly, accepted a job at Blackheath Private School, primarily due to her relationship with one of the teachers.

Regaining confidence – merging structure and inquiries

Laura's professional experiences, including those related to mathematics, change dramatically in her second year of teaching. After moving to Blackheath, she teaches grades 6-9, the levels she was educated for, and with whom she is more comfortable: "I don't want to comfort 3rd grade students, when they cry" [4]. There are fewer and less severe conflicts with the students, and the parents and the leadership are supportive. In contrast to her old school, "student learning is a joint venture between three parties: the teachers, the parents and the leadership" [4]. Finally, she now works with colleagues, who are closer to her own age, who are less focused on regulating student behaviour, and with whom she develops personal as well as professional relationships. Laura positions herself with confidence among them, and her ambition is to promote inquiry-based teaching: "This is my personal aim, to show that inquiry-based mathematics teaching is manageable" [2]. To pursue this aim, she two years later applies for and is appointed to the formal position of coach in mathematics.

It is remarkable how Laura regains her confidence with inquiry-based teaching at Blackheath. Her general experience of being a valued colleague and of belonging at the school is certainly conducive to that. The subject-specific elements are a result of

merging the world of regulation from Southbank with that of the reform-world, leading Laura to interpret the term of inquiry-based teaching differently and to reconsider the need to structure teaching-learning processes, also when students are involved in mathematical investigations. Both of these experiences relate to her involvement in the induction programme. Already in the individual lesson study at Southbank, Laura questions her prior understandings and says that she had "a wrong image of inquiry-based teaching. I thought it should be huge, wild … that we had to explore [real-world] things without using real maths" [2]. In contrast, she now says that it does not have to be investigations of real-world phenomena, it can also be pure mathematical investigations, but that the focus must be on the mathematics. Second, when planning the joint lesson study on the problem mentioned previously, Laura repeatedly asked for suggestions on how to structure the students' work "it's difficult to introduce the problem so I don't give them too much but still give them something" [2]. She panics before teaching the lesson in grade 6 at Blackheath, worrying that "it is going to be like before … chaos where nobody understands the objective and just does something" [2]. However, she succeeds in structuring students' work and her interactions with them, by using, among other things, two suggestions from the mentors: introduce the problem as a claim that the students shall confirm or disprove and visually represent and discuss students' ideas by using a chalk-drawn number board on the floor. In the programme, then, and at Blackheath, Laura develops a strong sense of professional belonging.

DISCUSSION AND CONCLUDING REMARKS

In this paper, we show how Laura's identities develop over the first 18 months of her career from struggling alone and losing her professional balance at Southbank to regaining professional confidence by merging structures and inquiries at Blackheath. We suggest that participation in the induction programme and her shift to Blackheath, which offered her a different context for her professional life, fuelled the changes.

Laura's first year in teaching turned out to be highly emotional, and her mainly negative feelings of helplessness and frustration at Southbank almost led her to leave the profession. These feelings were in particular fuelled by a conflict that resembles the one identified by Pillen et al. (2013) between teaching-learning processes favoured by herself and those promoted by her mentor and colleagues, the latter also supported by the students, their parents and the leadership. Having left her pre-service education, Laura felt uprooted and struggled alone for her reform-oriented ideas. In terms of identity, she lost her professional balance, as she experienced not being valued, not belonging to a professional community, and becoming worse, not better, at her job.

Laura, however, did not leave teaching, and in terms of the constellations of practice and career-decision that Cochran-Smith et al. (2012) found, she – at this stage – combined weak teaching with moving on. At Blackheath, Laura regained her professional confidence, and 18 months into her career, her teaching may be characterised as substantial in contents and rich in learning opportunities, that is, as relatively strong in Cochran-Smith et al.'s terms. She merged the world of regulation

from Southbank with that of the reform-world, which enabled her to structure the students' inquiry activities and interact differently with them. These developments were fuelled by the induction programme, and the programme as well as the new school constituted significant contexts for her experiences of becoming a valued colleague and of her professional belonging.

Laura appears, then, to be one of the teachers, who should be prevented "*from leaving the job for wrong reasons*" (Kelchtermans, 2017, p. 965, emphasis in original). The reasons that almost made her leave were contextual and highly dependent on emotional, social and cultural aspects of schools and classrooms, as were the ones that later made her stay. For instance, the Southbank students' harsh attitude to their teachers influenced Laura's contributions to classroom interactions and thus to the quality of the teaching-learning processes. Therefore, we suggest shifting the emphasis from *teacher* quality (cf. Cochran-Smith et al., 2012) to *teaching* quality as co-constituted by classroom interactions and by cultures at the school and beyond, if we are to understand identity as it relates to teachers decision to stay or leave. We argue that an experiential approach to identity allows us to develop such understandings.

REFERENCES

Blumer, H. (1969). *Symbolic interactionism. Perspective and method*. University of Los Angeles Press.

Charmaz, K. (2014). *Constructing grounded theory* (2nd ed.). Sage.

Cochran-Smith, M., McQuillan, P., Mitchell, K., Terrell, D. G., Barnatt, J., D'Souza, L., Jong, C., Shakman, K., Lam, K., & Gleeson, A. M. (2012). A longitudinal study of teaching practice and early career decisions: A cautionary tale. *American Educational Research Journal, 49*(5), 844-880. https://doi.org/10.3102/0002831211431006

Darragh, L. (2016). Identity research in mathematics education. *Educational Studies in Mathematics, 93*(1), 19-33. https://doi.org/10.1007/s10649-016-9696-5

Holland, D., Skinner, D., Lachicotte Jr, W., & Cain, C. (1998). *Identity and agency in cultural worlds*. Harvard University Press.

Hong, J. Y. (2010, 2010/11/01/). Pre-service and beginning teachers' professional identity and its relation to dropping out of the profession. *Teaching and Teacher Education, 26*(8), 1530-1543. https://doi.org/https://doi.org/10.1016/j.tate.2010.06.003

Jong, C. (2016). Linking reform-oriented experiences to teacher identity: The case of an elementary mathematics teacher. *The Journal of Educational Research, 109*(3), 296-310. https://doi.org/10.1080/00220671.2014.947398

Kelchtermans, G. (2017, 2017/11/17). 'Should I stay or should I go?': unpacking teacher attrition/retention as an educational issue. *Teachers and Teaching, 23*(8), 961-977. https://doi.org/10.1080/13540602.2017.1379793

Lave, J. (2019). *Learning and everyday life. Access, participation and changing practice*. Cambridge University Press.

Lutovac, S., & Kaasila, R. (2018, February 03). Future directions in research on mathematics-related teacher identity [journal article]. *International Journal of Science and Mathematics Education*(16), 759-776. https://doi.org/10.1007/s10763-017-9796-4

Pillen, M., Beijard, D., & den Brok, P. (2013). Professional identity tensions of beginning teachers. *Teachers and Teaching*, *19*(6), 660-678. https://doi.org/10.1080/13540602.2013.827455

Russ, R. S., Sherin, B. L., & Sherin, M. G. (2016). What constitutes teacher learning? In D. H. Gitomer & C. a. Bell (Eds.), *Handbook of research on teaching* (5th ed., pp. 391-438). AERA.

Skott, C.K., Falkenberg, L.L., & Honoré, I. R. (2021). New mathematics teachers' learning when participating in induction on mathematics education. A case study of two lower secondary teachers. *NOMAD*, 26(3-4), 71-90

Skott, J. (2019). Changing experiences of being, becoming, and belonging: teachers' professional identity revisited *ZDM - The International Journal on Mathematics Education*, *51*(3), 469-480. https://doi.org/10.1007/s11858-018-1008-3

Skott, J. (2022). Conceptualizing individual-context relationships in teaching: Developments in research on teachers' knowledge, beliefs and identity. [CERME12]. Twelfth Congress of the European Society for Research in Mathematics Education (CERME12), Bozen-Bolzano, Italy.

Wenger, E. (1998). *Communities of practice. Learning, meaning, and identity*. Cambridge University Press.

DIDACTIC SUITABILITY CRITERIA IN TEACHERS' PRACTICAL ARGUMENTATION IN THE PHASE OF DESIGN OF A LESSON STUDY CYCLE ABOUT FUNCTIONS

Telesforo Sol, Alicia Sánchez, Adriana Breda and Vicenç Font

University of Barcelona, Spain.

In this study eight teachers, who know the Didactic Suitability Criteria, participate in a cycle of Lesson Study and the aim is to identify the use of Didactic Suitability criteria in the practical argumentation of teachers during the design of a didactic unit of functions for students between 15 and 16 years old in compulsory education in Spain. The model proposed in Pragma-dialectics and the Toulmin's model are used to identify the practical arguments. Some uses of the criteria in the teachers' argumentation in the design are shown, in particular the epistemic and cognitive criteria.

INTRODUCTION

Some research on reflection conducted in Lesson Study (LS) experiences has identified the following phenomenon: when teachers reflect on their practice, they agree on criteria to guide it, which can be reinterpreted as Didactic Suitability Criteria (DSC), even when teachers are unaware of this theoretical construct (Hummes, 2022); but there is little research in which LS participants are previously aware of this construct.

In the framework of a LS experience in which participants know and use the DSC construct, our research question is the following: how do teachers use DSC in their argumentation to justify the group design of a class on functions? In accordance with this question, the objective of this paper is to analyze the use of DSC in the practical argumentation that supports the agreements that emerge in the design phase of a LS cycle on the topic of functions. To answer this question, we first identify episodes of practical argumentation (Lewiński, 2018) in the design of the functions lesson, and then, we analyze them using the ideal model considered in Pragma-dialectics (Eemeren & Grootendorst, 2003) and the Toulmin's (2003) model as a theoretical reference.

It is observed that some teachers go beyond using the criteria as a guideline for the design of a class, showing a broader reflection. Also, the epistemic and cognitive criteria are the most present criteria in the teachers' argumentation.

THEORICAL FRAMEWORK

Argumentation

We analyzed the argumentation of this study from the pragma–dialectical perspective (Eemeren & Grootendorst, 2003). Since pragma–dialectic proposes an ideal model for critical discussion, four stages can occur (or not) in this technique, namely: 1. *Confrontation stage*: Establishes the difference of opinion. In a non-mixed difference of opinion, this simply means that one party's point of view is not immediately accepted by the other, but instead is met with doubt or criticism. In a mixed difference

2023. 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. 227-234). PME 46.

of opinion, the other party advances its opposing point of view; 2. *Opening stage*: Refers to the starting points of the discussion and assigns the roles of protagonist and antagonist (in a mixed difference of opinion there are two protagonists and two antagonists). Moreover, the rules of the debate and the starting points are agreed upon; 3. *Argumentation stage*: The protagonist defends his/her point of view against the antagonist's persistent criticism, advancing arguments to meet the antagonist's objections or to remove the antagonist's doubts; 4. *Concluding stage*: The parties evaluate the extent to which the resolution of the difference of opinion is reached and in favor of whom. If the protagonist withdraws the point of view, the difference of opinion is resolved in favor of the antagonist; if the antagonist abandons his/her doubts, it is resolved in favor of the protagonist.

For the argumentation phase, since we are interested in knowing the use of DSC to justify actions that guide teaching practice, we consider practical argumentation, "argumentation aimed at deciding on a course of action" (Lewiński, 2018, p. 219). Gómez (2017) understands practical argumentation as that happening in social contexts that is oriented towards choosing an action to solve a practical problem. Practical argumentation answers questions such as 'what should a do in a situation x?', where a is the name or the description of an agent and x is the description of a problem situation. The argumentative model of Toulmin (2003) works this way: from some evidence (data), a claim is formulated. A warrant connects data with the claim, which is based on a theoretical, practical or experimental foundation: the backing. The modal qualifiers (surely, definitely, etc.) indicate how the claim is interpreted as true, possible, or probable. Its possible rebuttals or objections are also considered.

Didactic suitability

The didactic suitability of a teaching and learning process is defined as the degree to which such process (or a part of it) meets certain characteristics that allow it to be qualified as optimal or adequate to reach the adaptation between the personal meanings achieved by the students (learning) and the intended or implemented institutional meanings (teaching), considering the circumstances and available resources (environment). A teaching and learning process will achieve a high degree of didactic suitability if it is capable of articulating, in a coherent and systematic way, the following six partial criteria of didactic suitability, referring to each of the six dimensions involved in the teaching and learning process (Breda, et al., 2017): a) *Epistemic criterion*. To assess whether the mathematics that is taught is 'good mathematics'; b) *Cognitive criterion*. To assess, before starting the instructional process, whether what is intended to be taught is at a reasonable distance from what the students know; and after the process, what they have learned; c) *Interactional criterion*. To assess whether the interaction solves students' doubts and difficulties; d) *Mediational criterion*. To assess the adequacy of resources and time used in the instructional process; e) Affective criterion. To assess the students' involvement (interest, motivation) in the instructional process; f) *Ecological criterion*. To assess the

adaptation of the instructional process to the educational project of the school, the curricular guidelines, the conditions of the social and professional environment, etc.

Each didactic suitability criterion (DSC) has its respective components, and their utility requires defining a set of observable indicators that allow assessing the degree of suitability of each dimension of the teaching and learning process. **Error! Reference source not found.** presents the components of each DSC, based on Breda et al. (2017).

Suitability criterio	Components
Epistemic	Errors; Ambiguities; Richness of processes; Representativeness of the mathematical object complexity.
Cognitive	Prior knowledge; Curricular adaptation to individual differences; Learning; High cognitive demand.
Interactional	Teacher-student interaction; Student interaction; Autonomy; Formative assessment.
Mediational	Material resources; Number of students, class schedule, and classroom conditions; Time.
Affective	Interests and needs; Attitudes; Emotions.
Ecological	Curriculum adaptation; Intra and interdisciplinary connections; Social and labour usefulness; Didactic innovation.

Table 2: Didactic suitability criteria and components.

Breda et al. (2017) present rubrics with the components and indicators of each criterion. For instance, Table 2 is the rubric for the epistemic suitability.

Components	Indicators
Errors	Practices that are mathematically incorrect are not observed.
Ambiguities	Ambiguities that could confuse students are not observed; definitions and procedures are clearly expressed; explanations, evidence and demonstrations are suitable for the target level of education, the use of metaphors is controlled, etc.
Richness of processes	Relevant processes in mathematical activity (modelling, argumentation, problem solving, connections...) are considered in the sequence of tasks.
Representativeness of the complexity of the mathematical object to be taught	The partial meanings (constituted of definitions, properties, procedures, etc.) are representative samples of the complexity of the mathematical notion to be taught.
	For one or more partial meanings, a representative sample of problems is provided and a representative sample of different modes of expression (verbal, graphic, …) is provided.

Table 3: Components and indicators of epistemic suitability.

The DSC and their components are based on the principles and standards of the National Council of Teachers of Mathematics, trends in Mathematics Education and research in this area (Breda et al., 2017). Therefore, they constitute a consensual tool, which is used to structure teachers' reflection in teacher training programs in Spain, Chile, Brazil, Panama, Mexico, Argentina, and Ecuador (e.g. Hummes, 2022).

Lesson Study (LS)

The LS is the collaborative and detailed design of a lesson, its implementation and direct observation in the classroom, and its joint analysis after the implementation (Fernández & Yoshida, 2004). A group of teachers and experts with a common concern about their students' learning gather, plan a lesson and analyse and discuss what they observed in the implementation. A LS cycle should follow these phases: study of the curriculum and goals, when participants choose a content to teach and establish the learning goals; lesson planning, when participants set the objectives of the lesson and describe its development; implementation and observation of the lesson, when the impact of the planning on the students' learning is recorded and data generated from the observation are collected; joint reflection on the data collected, when participants use the data from the observation to reflect on the lesson implemented, the students' learning and the previous planning. For each phase of the cycle, some criteria should be considered in order to complete a LS cycle.

METHODOLOGY

This is a qualitative/interpretive research involving eight teachers (of mathematics and mathematics education) who are familiar with the DSC, with the purpose of designing a lesson on functions for 15-16 year old students of secondary school in Spain. The eight sessions of the first phase of the LS cycle (carried out so far) have been videotaped. In theme, the learning objectives, the number of sessions, some activities to work with students on the concept of function, previous knowledge, among other aspects, have been defined.

The analysis of argumentation is carried out considering the Pragma-dialectic model, the Toulmin's model and the characterization of practical argumentation. Following these phases: i) Review of videos to identify episodes of practical argumentation that show the use of DSC in the design of the class. ii) Identify the different ideas that teachers express when they participate. iii) Relate the ideas identified with the DSC. iv) Identify words that account for the existence of arguments, such as: then, therefore, for example, if, that is, etc.; and words that account for teachers' knowledge, values or beliefs: I believe, as far as I know, I understand that, we are supposed to, it is said, etc. v) Describe the dialogue to show the use of DSC in the design of the tasks for the class. vi) Apply the pragma-dialectic model. vii) In the argumentation stage, present the identified arguments, considering the Toulmin's model. viii) Describe the arguments taking into account the DSC.

The initial analysis of the arguments was performed by one researcher and subsequently triangulated by the rest of the researchers to clarify interpretations.

ANALYSIS

The following dialog is an example of how DSC are considered in the LS design phase of the didactic unit on functions:

T4: T6 says that we are making inputs [for design] without following the DSC.

T8: Yes, but we could make them according to the following criteria.

T1: I think that more or less. For example when we looked at the textbooks, we made a first analysis. Then we made some discussions based for example on the processes that are present in the textbook, on the meanings that are contemplated in the textbook. So just the fact of thinking about that I think that already comes from a knowledge we have of the DSC. For example, in the meetings we have had, in the discussions, I think they were. Well, many things were mixed.... many of the comments we made were related to the criteria, at least the first thing I looked at was to see which meanings were contemplated in the textbook and this is a knowledge that I have of the criteria, of the criterion of representativeness of the epistemic facet.

T8: Yes, but we can make comments that refine or improve some of these criteria, as something more than the use of a guideline. For example, I understand that we have looked at the different meanings that functions have and we talk about at least three meanings: Functions as a relation between quantities, first meaning, second, functions as a relation between variables and, third, a subset of the cartesian product. We concluded that we would start with a relation between magnitudes and we would arrive at a relation between variables, without marking well the difference between both meanings. That is, we would start with problems of relation between magnitudes, and, from a certain moment on, we would call these magnitudes variables, this is what I thought we agreed upon. So from the point of view of the criterion of the use of different meanings, there are at least two meanings. And maybe if we talk about the domain, we could make a little foray into the third meaning as a subset of the cartesian product. So we have made an analysis of the different meanings. It must be said that the current discourse on functions emphasizes the use of different representations, an indicator of the representational component of the meanings. Regarding the functions, it is always said that different representations must be worked, therefore we will put problems where the students have to handle tables, statements, graphs, that they have to find formulas. The subject of functions is taught this way nowadays, with the use of different representations. And then we were left with the indicator of using a variety of problems, ensuring that the problems are not always the same, and I would like to add an extension to this. For example, to relate it to cognitive suitability, it is assumed that we should try to make students learn the functions. Then there are all neuroscience studies that tell you that variation is important. This means that the student learns more if you vary than if you repeat, so they say that you have to vary the examples and you have to vary the tasks. So somehow here we would have a cognitive support, which would come in this case

> from neuroscience, about the fact of using different problems, that the problems are varied, that there is a variation of problems, we would have a support with the cognitive aspect. I don't know if you understand what I mean, but it is an example of how I go a little beyond what the guideline strictly says. [Three teachers agree with the comments among them T6.]

The dialogue presented is an example of how the DSC are part of the argumentation to justify the actions that are considered in the design of the different tasks for a class on the concept of function. In this particular case, teacher T4 stated that the contributions that were being made for the design of the class were not guided by the DSC. Professor T1 counterargued with two examples in which they were being used, in the first one she referred to the epistemic criterion (components richness of processes and representativeness of complexity), in the second one she commented that in the discussions many comments were made thinking about the DSC. Professor T8 presented an example of how the DSC are being used in the design. In his example, he mentions the epistemic criterion, keeping in mind the component of representativeness of complexity, considering three meanings for working in class, working with different representations and the consideration of a variety of problems. The latter is related to the cognitive criterion in its learning component.

Next, the above dialogue is presented from the perspective of Pragma-dialectics, and in the argumentation phase, the Toulmin's model is taken into account.

Confrontation: T6 states that DSC are not being considered as a guideline.

Opening: T1 and T8 consider that they are being considered.

Argumentation: *T1's arguments: Argument 1*: a) Data: Textbooks were analyzed; Discussions were made about the textbooks based on the processes that are present in the textbook and the meanings that are contemplated; b) Warrant: In the analyses and discussions, components of the DSC are considered; c) Backing: Understanding the DSC is to look at the meanings contemplated in a textbook; d) Claim: The DSC are being used in the design of the classes. *Argument 2*: a) Data: In the discussions, from the class design meetings, many things were mixed up; b) Warrant: Many of the comments were related to DSC; c) Claim: DSC are being used in class design.

To show the use of knowledge about the DSC in the analysis and discussion of the class design, T1 in her arguments exemplifies the use of the components of the epistemic criteria for the analysis of the textbook and its discussion. From this, it is inferred that a belief of T1 is that knowledge of the appropriateness criteria serves as a guideline in class design for analysis and discussion.

T8's arguments: Argument 1: a) Data: We have looked at different meanings of functions; from the DSC different meanings are being worked on; b) Warrant: Different meanings of mathematical objects should be taught; c) Claim: We would start with problems of relationship between magnitudes, and from a certain point on, these magnitudes will be called variables; d) Claim: We would start with problems of

relationship between magnitudes, and then these magnitudes will be called variables; e) Claim: We would start with problems of relationships between magnitudes.

In this practical argument, the "different meanings" component of the epistemic criterion serves to justify the order and the way of working different meanings of the function concept in class. This shows that T8 positively values designing a class considering the different meanings to be taught.

T8's arguments: Argument 2: a) Data: All discourse on functions considers the use of different representations; the different representations is an indicator of the representational component of the meanings; b) Warrant: We must work on different representations of the functions; c) Claim: We will set problems where the students will have to handle tables, statements, graphs and formulas.

In this practical argument T8 justifies working on the different representations of functions and relates it to the indicator "different representations" of the epistemic criterion and to the current trend in the educational discourse on functions.

T8's arguments: Argument 3: a) Data: The indicator of variety of problems is considered; We are supposed to try to make students learn the functions; b) Warrant: We must vary the examples and we must vary the tasks; c) Backing: There are neuroscience studies that say that the variation of examples and tasks is important; d) Claim: We must vary the examples and we must vary the tasks about functions.

To justify the variation of examples and tasks on functions, in this practical argument T8 uses two DSC. First, the epistemic criterion with its indicator "variety of problems" and, second, the cognitive criterion with its indicator "learning", supporting the relationship between these two indicators with results from neuroscience.

DISCUSSION AND CONCLUSIONS

This work has the particularity that the teachers participating in the LS cycle are aware of the DSC, unlike previous research where most or all of them are unaware of them. Moreover, an increase in the time dedicated to the design phase has been observed with respect to the LS in which the participants are not aware of the DSC. A plausible explanation, although not the only possible one, is that knowledge of the DSC allows teachers to broaden the design discussion and structure arguments, which contain the DSC, to justify the actions to be considered.

The research objective is to analyze the use of DSC in the practical argumentation that occurs in the design phase of a LS cycle on the topic of functions. The following uses of DSC were identified: a) their application for the analysis of textbooks, in particular the processes and meanings contemplated by the book on the teaching of a mathematical object; b) to discuss about the design of the class, in particular they are used to generate guarantees to justify the proposed actions, and c) they are used as data in the practical arguments.

While there are arguments in which only one DSC is used, in others, several are used and related to justify the proposed actions, which we could interpret as stronger arguments because they are more difficult to attack in the discussion.

The analysis of the videos shows that the most used criterion for the design of the class on functions is the epistemic criterion, followed by the cognitive criterion. It is also observed how the DSC are particularized for the design of the class on the notion of function (for example, the component that states that a variety of meanings of the mathematical object must be taken into account is specified in three meanings of the notion of function: relation between magnitudes, between variables and subset of the cartesian product).

ACKNOWLEDGEMENTS

PID2021-127104NB-I00 (MICINN,FEDER, UE) and grant PRE2019-090177 (funded by Ministerio de Ciencia e Innovación).

REFERENCES

Breda, A., Pino-Fan, L. R., & Font, V. (2017). Meta Didactic-Mathematical Knowledge of Teachers: Criteria for The Reflection and Assessment on Teaching Practice. *Eurasia Journal of Mathematics, Science and Technology Education, 13*(6), 1893-1918. https://doi.org/10.12973/eurasia.2017.01207a

Eemeren, F., & Grootendorst, R. (2003). *A Systematic Theory of Argumentation: The pragma-dialectical approach*. Cambridge University Press. https://doi.org/10.1017/CBO9780511616389

Fernández, C. & Yoshida, M. (2004). *Lesson study: a Japanese approach to improving mathematics teaching and learning.* Erlbaum.

Gómez, J. (2017). ¿Qué es la argumentación práctica? [What is practical argumentation?]. *Revista Co-herencia, 14*(27), 215–243. https://doi.org/10.17230/co-herencia.14.27.9

Hummes, V. (2022). *Uso combinado del Lesson Study y de los Criterios de Idoneidad Didáctica para el desarrollo de la reflexión sobre la práctica en la formación de profesores de matemáticas* [Doctoral Dissertation, University of Barcelona]. UB Repository. http://hdl.handle.net/10803/675835

Lewiński, M. (2018). Practical argumentation in the making: Discursive construction of reasons for action. In S. Oswald, T. Herman, & J. Jacquin (Eds.) *Argumentation and Language – Linguistic, Cognitive and Discursive Explorations* (pp. 219–241). https://doi.org/10.1007/978-3-319-73972-4_10

Toulmin, S. (2003). *The Uses of Argument* (2nd ed.). Cambridge University Press. (Original work published 1954).

FOSTERING STUDENTS' KNOWLEDGE ABOUT PROOF

Femke Sporn, Daniel Sommerhoff and Aiso Heinze

IPN - Leibniz Institute for Science and Mathematics Education, Kiel, Germany

Empirical research has underlined difficulties regarding knowledge about proof for learners of all ages. As knowledge about proof can indicate an individual's understanding of proof, it is thus mandatory to find means to support students effectively. In an intervention study with 61 9th grade students, we examined to what extent it is possible to foster students' understanding of proof by explicitly addressing knowledge about proof principles. The intervention covered five lessons integrated in regular mathematics classes distributed over a half school year. The positive effect suggests that the intervention could foster students' knowledge about proof principles using only a short amount of time. However, the results also raise questions regarding the explicit formulation of rules and criteria for mathematical proof.

INTRODUCTION

In their mathematics education, learners are repeatedly confronted with argumentative challenges, mathematical proofs, and related activities. Thus, it can be assumed that learners (are expected to) form an understanding of proof throughout their schooling, which can be understood as an understanding of the way evidence is generated in mathematics (Sporn et al., 2022). Individual's understanding of proof has various aspects and, for example, includes learners' knowledge about proof and their need for proof. Building an adequate understanding of proof is relevant for learners (i) because they are expected to learn about mathematics as a deductive system and how evidence is generated in mathematics as well as (ii) because there is repeated evidence that an adequate individual's understanding of proof improves performance in handling proof (e.g., Chinnappan et al., 2012). Still, prior research indicates that learners of different ages show an inadequate understanding of proof (e.g., Healy & Hoyles, 2000).

Currently, there is still insufficient knowledge on how an individual's understanding of proof (or single aspects of it) is formed throughout mathematics education, which factors influence its formation, and how it can be explicitly supported. It seems likely that the learning opportunities regarding mathematical proof throughout mathematics education and regarding various proof activities, such as validating or constructing mathematical proofs, play a role in building an individual's understanding of proof. Opportunities for reflection on and discussions about proof in mathematics are also possible, likely especially useful learning opportunities (e.g. Davis, 2000). However, there is insufficient information on such learning opportunities and their impact so far.

Since learners repeatedly show difficulties in their proof performance and an inadequate understanding of proof, our focus was on designing and evaluating a teaching intervention aiming to foster 9th grade students' understanding of proof. We focused on students' knowledge about proof (as one aspect of their understanding of

2023. 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. 235-242). PME 46.

proof) to investigate to what extent such an intervention is indeed effective in fostering students' knowledge about proof and thus their understanding of proof.

THEORETICAL BACKGROUND

Learning Opportunities Regarding Mathematical Proof in Secondary Education

Reasoning and proving are integrated in mathematics curricula worldwide (e.g., CCSSI, 2010). Often, proofs are explicitly focused on in geometry classes from grades 7 and 8 onwards. Thus, learners repeatedly encounter learning opportunities for mathematical proof. While there is a lack of empirical evidence if and to what extent those learning opportunities impact on learner's understanding of proof in specific, it can be assumed that the formation of an individual's understanding of proof is positively affected by such (different) learning opportunities regarding mathematical proof (Sporn et al., 2022).

An objective evaluation of such learning opportunities and, thus, also of the (learning) output is only possible to a limited extent since a generally accepted definition of a valid mathematical proof is missing in mathematical practice (Stylianides et al., 2017). In particular, the social context (e.g., a course) and the norms and criteria established therein – possibly initiated by the teacher and oriented to the mathematics standards – are of a certain influence (Sommerhoff & Ufer, 2019). Thus, a specific argumentation might be considered a valid mathematical proof in one context and, at the same time, be rejected in another context (e.g., Inglis et al., 2013).

Students' Knowledge About Proof (Principles) as one Aspect of Their Understanding of Proof

While there are certain social aspects related to the norms and criteria for mathematical proof, there are also many aspects of mathematical proof that are widely agreed upon and which learners should learn, including also knowledge about the way evidence is generated in mathematics (i.e., *knowledge about proof*). This includes criteria that must be fulfilled for a valid mathematical proof independently of the social context (e.g., the argumentation must be based on already proven statements). Such criteria, that a mathematical proof must definitely fulfill, can be summarized as *proof principles*.

As part of the formation of an individual's understanding of proof, learners build up *knowledge about proof principles*, which has been addressed, for example, by Heinze and Reiss' (2003) research on proof scheme, proof structure, and chain of conclusions, and which has shown that 8th grade learners have difficulties regarding proof principles. The findings of Sporn et al. (2022) suggest that students' knowledge about proof principles may develop slightly positively over the course of mathematics education. However, it is not yet known, which kinds of learning opportunities explicitly allow to foster students' knowledge about proof principles, and thus to what extent these allow to specifically support students' development of understanding of proof.

Prior research shows two different ways of operationalization (Sporn et al., 2021): (i) knowledge about proof principles in an explicit and general form, that is without

reference to a specific proof-related action situation (*concept-oriented knowledge about proof principles*; e.g., Heinze and Reiss (2003)), and (ii) knowledge about proof principles with a focus on a specific mathematical action situation regarding proof, for example, the construction or validation of a specific proof (*action-oriented knowledge about proof principles*; e.g., Andersen (2018)). That a distinction between the two foci is theoretically and also empirically meaningful for investigating knowledge about proof principles was shown by Sporn et al. (2021). Figure 5 shows an example item for an action-oriented focus on knowledge about proof principles. Here, students are required to validate a purported proof that does not follow the established proof principles (i.e., proof by authority).

Ben has to prove the following proposition:

The sum of three consecutive natural numbers is divisible by 3.

Ben's purported proof:

I know this from school. Our textbook contained a proof that this is valid for every natural number. There, it was shown that:

$$3 + 4 + 5 = 3 + 3 + 1 + 3 + 2 = 3 \cdot 3 + 3$$

This proves the proposition.

Is Ben's purported proof a valid mathematical proof? ○ **Yes** ○ **No** ○ **Don't know**

Figure 5: Example item used in the study for measuring action-oriented knowledge about proof principles (translated; Sporn et al., 2022).

In comparison, evaluating the statement *"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."* addresses concept-oriented knowledge of proof principles.

RESEARCH QUESTIONS

Since (i) prior research has highlighted students difficulties regarding their knowledge about proof principles (e.g., Healy & Hoyles, 2000) and that the desired development of students' knowledge about proof principles in school could not be verified (Sporn et al., 2022) and (ii) research and information on whether, how exactly, and how effectively their knowledge about proof principles can be fostered is missing, the question arises whether it is possible to explicitly foster students' (concept- and action-oriented) knowledge about proof principles. To examine this question, learning opportunities regarding mathematical proof were developed that induce, for example, explicit discussion and reflection on proof principles. In the context of an intervention, the aim was to investigate whether these learning opportunities have a positive effect on students' (concept- and action-oriented) knowledge about proof principles as well as their justifications regarding the acceptability of proofs. The research questions thus focused on: (RQ1) How does an intervention on mathematical proof and proof principles affect the concept- and action-oriented knowledge about proof principles of 9th grade students? (RQ2) How does an intervention on mathematical proof and proof principles foster 9th grade students' ability to use criteria aligned with proof principles to justify their judgments of incorrect purported proofs?

Regarding RQ1, it was expected that discussing and reflecting on proof principles during the intervention would lead to an explicit formulation of proof principles, thus

directly addressing a concept-oriented knowledge about proof principles and thus causing increased knowledge (H1.1). By elaborating and explicitly reflecting on these criteria while working on specific proofs, it was expected that action-oriented knowledge about proof principles would improve (H1.2). Regarding RQ2, students were expected to judge purported proofs using criteria aligned with proof principles more often (H2), as these were explicitly introduced to students in the intervention.

METHOD

Study Design & Sample

In total, 61 students from four 9th grade classes of a German Gymnasium participated in a quasi-experimental intervention study with an intervention and a control group. A questionnaire was used to assess their concept- and action-oriented knowledge about proof principles at the beginning of the second half of 9th grade (pre-test). Two of the four classes received an intervention on mathematical proof (intervention group; $N_{\text{Int}} = 30$). Both other classes received no specific intervention (control group; $N_{\text{Con}} = 31$). At the end of the second half of the 9th grade, students' concept- and action-oriented knowledge about proof principles was assessed (post-test), again.

Design of the Intervention on Mathematical Proof

The intervention included five lessons (one double lesson (T1) and three single lessons (T2-T4); each of which included additional homework), which were distributed over the second half of the school year with in-between intervals of about 3-4 weeks. The intervention focused, for example, on discussing and reflecting on proof principles based on the treatment of different (purported) proofs and thus focused on students' knowledge about proof principles. Proofs for mathematical statements were selected, which are part of the 9th grade curriculum anyway. At T1, the focus was on the role of proof in mathematics and on the distinction between premises and conclusion. Students were supposed to become familiar with proofs (as objects - cf. Reid & Knipping, 2010) quite consciously and their need for proof was encouraged. Problematization took place at the end of T1: students developed the central question whether and how to decide when a purported proof can be judged as a valid or invalid mathematical proof. Answering this question was the focus of the three single lessons T2-T4. At T2 and T3, the focus was on reflecting on proof principles (based on proof scheme, proof structure, and chain of conclusions; Heinze and Reiss (2003)). Students used purported proofs to elaborate in which cases one can be sure to say that a purported proof is invalid (i.e., when proof principles were not followed) and then abstracted the proof principles in explicit and general terms (i.e., argumentation should not be based on experience; using only single examples to show an assertion is not sufficient, etc.). Like this, students elaborated on the proof principles (action-oriented focus on knowledge about proof principles) and formulated them in explicit and general terms (concept-oriented focus on knowledge about proof principles) (knowledge integration, see Linn et al., 2013). Following up on the problematization from T1, the aim at T4 was to clarify if or when it can be decided that a purported proof represents a valid

mathematical proof. For this, the insights on proof principles from T2 and T3 were integrated by stating that while there is no way of judging the acceptability of a proof absolutely, there is at least no reason to judge a purported proof as invalid if all proof principles are followed (e.g., Davis, 2000). Again, students used purported proofs to discuss and reflect on proof principles with the aim of finding a general formulation on the fact that it finally depends on the social context when a purported proof is judged as a valid mathematical (i.e., depending on prior knowledge, one course may require more intermediate steps in the proof than another course to accept it; Yackel and Cobb (1996)). At the end of each lesson, the results were saved on the whiteboard and sent to the project team. The results were checked for correctness, and the beginning of the following intervention lesson was slightly adapted to these results (for example the wording of the respective class is used; e.g., *"The purported proof is invalid, if it is based on the statement of a book."*). Results were also saved on a poster in the classroom to have the elaborated explicit formulations available at any time and to repeat them before continuing with the new content. The poster was visible to the students throughout the school year (except at the time of the post-test). The teachers of the intervention group were provided with the same materials and received very detailed instructions on how to implement each lesson. The control group did not receive any additional materials but followed the regular mathematics curriculum.

Items & Analyses

The employed questionnaires were based on prior studies (see Sporn et al., 2022). At the beginning and at the end of the intervention, an (identical) questionnaire was used, which included 18 statements on concept-oriented knowledge about proof principles. These included valid and invalid statements about proof principles, which had to be evaluated on a 6-point Likert scale ("Not true at all" (1) to "Totally true" (6)). The 18 statements were combined to a mean score S_{coK} and rescaled to a range from 0 to 6 for better comparability. Action-oriented knowledge about proof principles was assessed by asking students to validate six purported proofs in each of which certain proof principles were disregarded. Figure 5 shows an item example. For each judgment, students scored 1 point if they correctly judged the purported proof to be invalid. In each other case, they scored 0 points. A sum score was formed from the judgments of the six purported proofs (S_{aoK}; range: 0 to 6). In addition, students were asked to provide a justification for their judgment about the validity in an open text box for each purported proof. For the justification of their judgment, students scored 1 point if they justified the correct judgment (i.e., *invalid purported proof*) using criteria aligned with proof principles. They scored 0 points if they justified their judgment in a way that was not aligned with proof principles or if they made an incorrect judgment (i.e., *valid purported proof*). Again, a sum score was formed ($S_{\text{aoK-Justification}}$; range: 0 to 6). To answer the research questions, ANCOVAs were calculated for each of the three scores (S_{coK}, S_{aoK}, $S_{\text{aoK-Justification}}$).

RESULTS

Figure 6 shows the descriptive results of the intervention and control group regarding the three scores (S_{coK}, S_{aoK}, $S_{\text{aoK-Justification}}$) in the pre- and post-test. The graphs indicate cross-over interactions for each score, implying that the intervention group achieved better learning results than the control group.

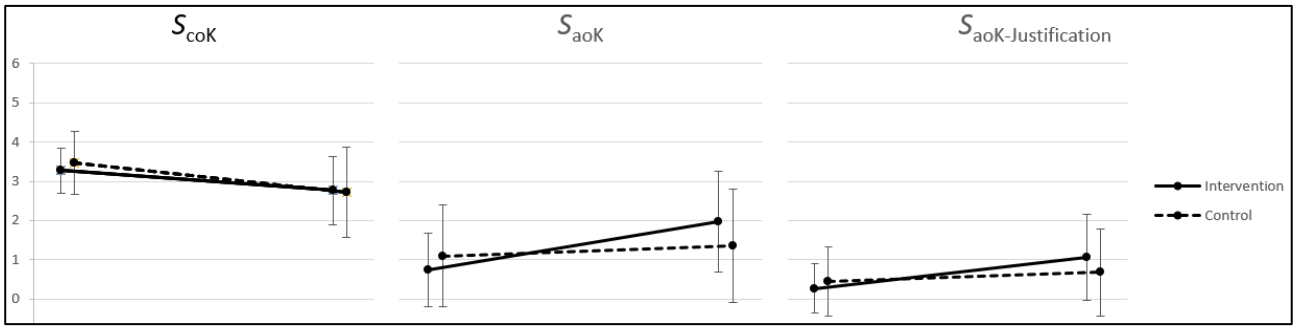


Figure 6: Descriptive results (Mean, SD) for S_{coK}, S_{aoK}, $S_{\text{aoK-Justification}}$.

Descriptive data for S_{coK} show a (non-significant) decrease for both groups. The ANCOVA for S_{coK} shows that after adjusting for the pre-test, post-test S_{coK} does not differ statistically significantly between both groups ($F(1, 59) = 0.11$, $p = .744$, $\eta_p^2 = .002$). Although the students were encouraged to explicitly formulate proof principles as part of the intervention, this does not seem to have impacted their concept-oriented knowledge about proof principles. While S_{aoK} was low in the pre- and post-test, the intervention appears to have a positive effect, at least for some students. The ANCOVA for S_{aoK} shows a significant difference between both groups ($F(1, 58) = 5.82$, $p = .019$) in the post-test with a medium to large effect ($\eta_p^2 = .091$) in favor of the intervention group. The ANCOVA for $S_{\text{aoK-Justification}}$ shows similar results: again, the intervention group scored significantly higher than the control group ($F(1, 58) = 5.91$, $p = .018$, $\eta_p^2 = .092$) in the post-test. Thus, with intervention, students were more likely to evaluate the incorrect purported proofs correctly as invalid and were also more likely to do so based on criteria aligned with proof principles.

DISCUSSION & OUTLOOK

Prior research has repeatedly shown that learners have shortcomings regarding multiple aspects of their understanding of proof (e.g., Chinnappan et al., 2012), for example regarding their knowledge about proof (Heinze & Reiss, 2003). It thus appears essential to find ways to effectively foster such knowledge about proof and by this students' understanding of proof. As information about how to foster students' knowledge about proof (principles) is currently missing, an intervention on mathematical proof (principles) was designed and evaluated. By using specific mathematical action situations regarding proof to elaborate on proof principles (action-oriented focus on knowledge about proof principles) and by developing formulations of proof principles in explicit and general terms (concept-oriented knowledge about proof principles), 9th grade students were encouraged to discuss and reflect on proofs (Davis, 2000; Linn et al., 2013).

Contrary to expectations, empirical results show no significant difference between both groups regarding their concept-oriented knowledge about proof principles in the post-test (H1.1). As elaborating and formulating proof principles in explicit and general terms appears to be a reasonable approach to support concept-oriented knowledge about proof principles, it remains unclear why no effects could be observed. For further insights, either qualitative approaches would be needed or longitudinal studies with more measurement points, which would allow observing direct and cross-lagged effects between time points to better understand the longitudinal relations between gains action- and concept-oriented knowledge. Moreover, the results may also indicate measurement issues regarding concept-oriented knowledge, as students were possibly not able to see the equivalence of the formulations of proof principles in the tests and their formulations in classroom. As students created their own formulations for proof principles (verified as correct) in class, they may not have recognized that they corresponded to the statements presented in the questionnaire. Thus, such an item format may not be useful for assessing concept-oriented knowledge about proof principles specified in the context of the intervention. Prompting learners for proof principles in an open item may be more appropriate (Andersen, 2018). Even though action-oriented knowledge about proof principles is low in the pre- and post-test (corresponding to results of prior research, e.g., Heinze & Reiss, 2003), a ANCOVA shows a significant positive effect of the intervention on the action-oriented knowledge about proof principles, as expected (H1.2). Thus, even a relatively short intervention (albeit over a longer period of time) including, for example, some encouraging opportunities for discussion and reflection on proof principles, seems to have a medium to large effect on students' action-oriented knowledge about proof principles. This suggests that fostering students' knowledge about proof is possible within half of a school year. Regarding RQ2, as expected, the intervention students are significantly better at using criteria aligned with proof principles to justify their judgments of incorrect purported proofs. This supports the results on the significant positive effect of the intervention, and shows that students from the intervention group are more likely to be able to formulate criteria aligned with proof principles. This result may also indicate that they learned the proof principles, even if not in the form of the statements presented in the questionnaire for concept-oriented knowledge about proof principles.

As with every research, the presented study has some limitations, for example regarding the number of participants and the restriction to one school in one country. Further, it was not possible to control instruction in the control group. Albeit these limitations, empirical data suggests that it is useful for students to explicitly elaborate proof principles within discussion and reflection using different (purported) proofs and within various proof activities. Further research is necessary to examine, if other learning opportunities might lead to (even higher) effects on students' knowledge about proof and which modifications (length, arrangement of the sessions, methods, etc.) of the current intervention could be made to also foster concept-oriented knowledge. Furthermore, it would be interesting to investigate, to what extent such an increase in knowledge has an impact on students' performance in handling proof.

REFERENCES

Andersen, L. E. (2018). Acceptable gaps in mathematical proofs. *Synthese*, *197*(1), 233-247.

CCSSI. (2010). *Common core state standards for mathematics*. National Governors Association Center for Best Practices and the Council of Chief State School Officers.

Chinnappan, M., Ekanayake, M. B., & Brown, C. (2012). Knowledge use in the Construction of Geometry Proof by Sri Lankan Students. *International Journal of Science and Mathematics Education*, *10*(4), 865-887.

Davis, E. A. (2000). Scaffolding students' knowledge integration: prompts for reflection in KIE. *International Journal of Science Education*, *22*(8), 819-837.

Healy, L., & Hoyles, C. (2000). A Study of Proof Conceptions in Algebra. *Journal for Research in Mathematics Education*, *21*(4), 396.

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.

Inglis, M., Mejia-Ramos, J. P., Weber, K., & Alcock, L. (2013). On Mathematicians' Different Standards When Evaluating Elementary Proofs. *Topics in Cognitive Science*, *5*(2), 270-282.

Linn, M. C., Eylon, B.-S., & Davis, E. A. (2013). The knowledge integration perspective on learning. In M. C. Linn, E. A. Davis, & P. Bell (Eds.), *Internet environments for science education* (pp. 57-74). Mahwah, NJ: Lawrence Erlbaum Associates.

Reid, D. A., & Knipping, C. (2010). *Proof in mathematics education. Research, learning and teaching*. Rotterdam, Netherlands: Sense Publishers.

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*, *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.

Sporn, F., Sommerhoff, D., & Heinze, A. (2022). Students' Knowledge About Proof and Handling Proof. In C. Fernández, S. Llinares, Á. Gutiérrez, & N. Planas (Eds.), *Proceedings of the 45th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 27-34). Alicante, Spain: PME.

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.

FROM 2D TO 3D: SUPPORT OF A 3-DIMENSIONAL DYNAMIC GEOMETRY ENVIRONMENT IN LEARNING PROOF

Camilo Sua, Angel Gutiérrez and Adela Jaime

Universidad de Valencia, Spain

Dynamic geometry environments support the learning of proof in plane geometry. Researchers have studied this process by using theoretical frameworks that allow us to understand how these environments provide such support. However, there is scarce research carried out on 3-dimensional dynamic geometry environments, which rise questions about the ways in which the learning of proof occurs in this scenario and how those environments intervene in the process. Based on the case of a mathematically gifted student, we analyze the development of the student's proving skills while solving a sequence of construction-and-proof problems in a 3-dimensional dynamic geometry environment and the way in which the environment stimulated those skills through utilization schemes put to work by the student to use some tools.

INTRODUCTION

Research on the influence of dynamic geometry environments (hereafter, DGE) is notable and has a long history. One of the several aspects of teaching and learning geometry with DGE on which research has been developed in recent years (Sinclair et al., 2016) is learning of proof (Sinclair & Robutti, 2013). However, this research has been carried out mainly in two-dimensional DGE (2D-DGE) and there is a lack of related research on three-dimensional DGE (3D-DGE) (Gutiérrez & Jaime, 2015). Particularly, it's necessary to carry out research informing on the influence that 3D-DGE can have in the learning, by ordinary and mathematically gifted students, of spatial geometry and, in this mathematical context, the learning of mathematical proof.

Research on the influence of DGE in the learning of mathematics has used different theoretical frameworks to interpret this process through the actions that individuals perform when using a technological artifact (mainly software in a computer, tablet, etc.). Some of these frameworks are based on the premise that, through the actions of a person with an artifact, the cognitive activity that occurs in that person's mind can be understood, so it is possible to provide observable evidence about the mental processes performed by the students (Drijvers et al., 2009).

We have carried out a case study research where four mathematically gifted students solved a sequence of construction-and-proof problems in a 3D-DGE based on GeoGebra (i.e., problems asking to create a geometrical figure and then prove that the figure fulfils the conditions required by the problem). The objective of that research was to analyze the students' reasoning processes and their progress in learning to do deductive proofs, to get information about mathematically gifted students' learning trajectories and different styles of mathematical reasoning. This paper focuses on one of those students, with the research objectives of i) analyzing the improvement of his

2023. 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. 243-250). PME 46.

proving skills when he solved the mentioned sequence of problems and ii) showing the way the 3D-DGE helped promote this change, through the utilization schemes put into action by the student when he used some GeoGebra tools.

THORETICAL BACKGROUND

Construction-and-proof problems and the learning of proof

DGEs support the learning of proof (Mariotti, 2012). We consider a *proof* as a mathematical argument, either empirical or deductive, aimed to convince someone of the truth of a mathematical statement (Fiallo & Gutiérrez, 2017). In our study, we emphasize the learning of proof through *construction-and-proof problems*. These problems ask i) to create on the DGE a geometric figure having some properties required by the problem, that must be preserved under dragging, and ii) to prove that the procedure used to create the figure is correct, by explaining and validating the way of construction (Mariotti, 2019). The statement to be proved is that the sequence of actions of the construction fits the conditions of the problem.

By using DGE tools in the construction of geometric objects, personal meanings are produced thanks to the dependency interrelationships that are discovered and verified through dragging. These tools are also related to theoretical elements of the Euclidean geometry which can support students when they develop proofs for the constructions (Mariotti, 2012). Solving construction-and-proof problems allows students to take advantage of the DGE possibilities and the logical system that underlies it. Therefore, geometric constructions have also a purely theoretical nature, where their validity is linked to prove that a set of constructions steps provide a specific result, so solving this kind of problems can make students evoke theoretical meanings of the tools they have used in the solutions (Mariotti, 2019).

From artifacts to instruments: instrumentalization and instrumentation

An *artifact* is any object used as a tool to perform a task (Rabardel, 1995). When a subject establishes a relationship with an artifact to do a specific task, in which the artifact is used in a particular way for a specific purpose, the artifact becomes an instrument. An *instrument* is a theoretical notion, the combination of an artifact and some mental schemes developed by the user, to which they refer when using the artifact to perform a task (Rabardel, 1995). A *scheme* is an invariant organization of mental habits for a group of situations, a stable way of dealing with specific tasks (Vergnaud, 1996). For Rabardel, the transition from an artifact to an instrument requires two intertwined processes that come from the individual's relationship with the artifact: *instrumentalization*, seen as the recognition of the components of an artifact, its limitations, and possibilities to solve a task; and *instrumentation*, seen as the emergence and development of *utilization schemes* on the artifact when solving tasks.

Due to their nature, the schemes are not directly observable, so it is necessary to have an observable counterpart to be able to refer to them. We describe the schemes in terms

of students' behavior while are using the artifacts provided by a DGE, like dragging, construction tools, etc.

METHODOLOGY

We present a case study drawn from a broader research project where we analyze the learning of proof, in the context of a GeoGebra 3D-DGE, by four Spanish mathematically gifted students (11 to 14 years old) in grades 1 to 4 of secondary school. The identification of the students as mathematically gifted is because, for several years before our experiments, the students had participated in special out-of-school programs of attention to generally gifted students (AVAST) and mathematically gifted students (ESTALMAT) where they attended mathematics workshops.

We designed a sequence of 18 construction-and-proof problems which involved the equidistance relationship between points and between points and lines. Some problems requested the construction of a geometric object satisfying certain properties associated with equidistance (e.g., construct an equilateral triangle), first in 2D and then in 3D. Other problems requested first the construction of a 2D object and then the construction of an analogous object in 3D (e.g., construct the center of a circle in 2D and a sphere in 3D). These problems were implemented in several 60-minute sessions. The students' solutions of each problem provided instrumental and conceptual elements useful to solve subsequent problems. Students solved each problem and then discussed their solution with the first author of the paper, who led the conversation to justify the results. These sessions were audio and video recorded. As students were in different school grades and had different previous knowledge, the experimental sessions were organized as individual clinical interviews.

In this paper we present episodes of the solutions of three problems by a student named Hector (pseudonym). We chose those episodes because they show the development of Hector's skills for the elaboration of proofs and how he benefited from utilization schemes of some GeoGebra tools he created.

Table 1: Indicators of instrumentalization

Code	Indicator	Description
Tsa1	Discover possibilities of a tool	Previously unknown possibilities and functions of a tool (or set of tools) that allow solve the task are discovered
Tsa2	Identify limitations of a tool	An inadequate result is identified when using a tool with a defined purpose and such a purpose is ruled out for the tool.
Tsa3	Customize and adjust a tool to personal interests	Various uses of a tool are identified, according to specific interests. The tool is used in different schemes according to the requirements of the task.
Tsa4	Appropriate the artefact	The user becomes aware that the artifact is useful to obtain a specific result in a particular context.

Code	Indicator	Description
Sas1	Identify a scheme associated with one or more tools	A set of steps with one or more tools is recognized as effective to obtain a particular result. The set of steps was not known before.
Sas2	Elaborate a scheme to obtain a particular result	A scheme is elaborated/defined when using the tool that leads to obtain the same result.
Sas3	Adapt a scheme when solving a problem	A scheme is improved by including or removing some steps, to give it more scope or refine it.
Sas4	Use the same scheme in different tasks	The scheme that has been elaborated/modified is routinely used when solving different problems.

Table 2: Indicators of instrumentation

To analyze Hector's instrumental activity, we use an adaptation of the indicators of instrumentation and instrumentalization proposed by Sua and Camargo (2019) (Tables 1 and 2), which offer an analytical device to characterize both processes and come from the interpretation of their corresponding definitions in the specialized literature. The indicators of each process suggest a possible trajectory followed by a subject who uses artifacts to solve different tasks, leading them to become progressively instruments.

THE CASE OF HECTOR: THE CONSTRUCTION OF A BISECTOR PLANE

The episodes presented below report the evolution of the drawing of the perpendicular bisector of a segment with ruler and compass, which was transformed into a procedure with 3D-DGE that allowed the construction of the bisector plane of a segment. Solving previous problems had allowed introducing the definition of circle, sphere, and bisector plane as locus, as well as the definition of perpendicular bisector as a perpendicular line to the segment through its midpoint, the equidistance property of this line and that the bisector plane of a segment contained its perpendicular bisectors and therefore each point in this plane was equidistant from the ends of the segment. Problems that we present below used these properties and definitions. Before solving the problems, Hector already knew the definitions of sphere, circle, and perpendicular bisector. The other properties and definitions were obtained by solving previous problems.

Problem 7: Construction of an equilateral triangle in 2D and 3D

The first part of this problem asked to construct in GeoGebra 2D an equilateral triangle having the given segment GH as a side. To construct this triangle, Hector created the circles with centers on G and H and radius GH (Figure 1a).

To prove the validity of the construction, Hector expressed that he did not know why it worked, but that was the way he had been taught at school. However, the conversation with the teacher led him to look for support in the congruence of the radii of the circles, to state that the sides of the triangle were congruent because: ... *as these*

two circles are equal... as we have created the circles, so that the radius is this [pointing at GH], *these two sides* [IG and IH] *are the same… so, they are all the same.*

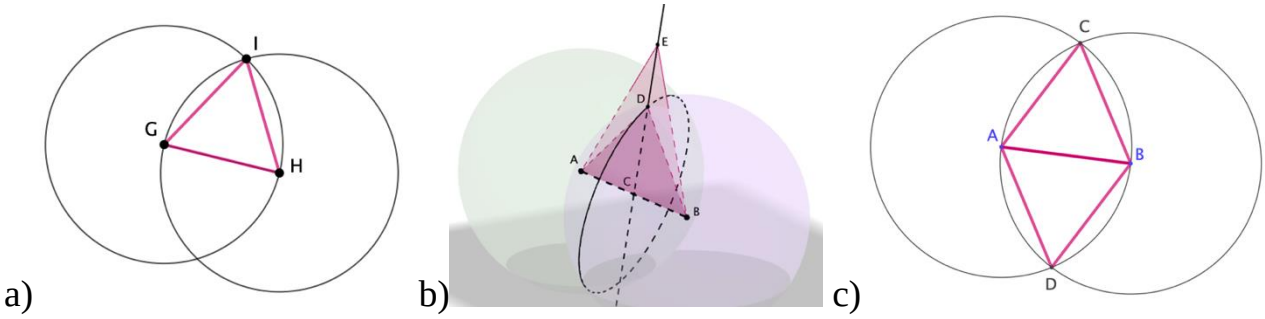


Figure 1. Construction of an equilateral triangle in 2D and 3D

The second part of the problem asked to construct an equilateral triangle in GeoGebra 3D having the given segment AB as a side. Hector created the midpoint C of AB, the spheres with centers on A and B and radius AB, and a point D at the intersection of these and the line CD (Figure 1b). Hector first described the construction: ... *I have done the same as in the 2D version, only that instead of circles I have used spheres ... I learned the process that I used before for the equilateral triangle to make the perpendicular bisector with a compass on the subject of arts ... I have used that same technique to make the perpendicular bisector*.

Hector proved the result of his construction considering that *the point* [D] *is on the perpendicular bisector... that is, these two sides* [AD and BD] *are equal. I have used spheres of radius AB, so these two segments here* [pointing to AD and BD] *are radius of circles* [he means spheres] *with radius AB. So, they're all the same*.

Although Hector had built the requested triangles, at the end the professor talked to him about the mechanism with ruler and compass that had been used by Hector, but it was not known why it worked. The objective of this conversation was to provide an explanation about that mechanism, given its relevant role in the actions carried out by the student. Hector made the construction showed in Figure 1c and, although it was not easy to him at first, he proved that line CD is perpendicular to segment AB, by using the fact that ADBC is a rhombus, because its sides are congruent, so its diagonals bisect each other and are perpendicular.

Problem 10: Construction of the bisector plane of segment AB

The objective of this problem was to discover that, given three points A, B and C, the intersection of the bisector planes of A, B and A, C was contained in the bisector plane of B and C. Therefore, it was necessary to construct the first two mentioned planes and drag point A to different places in space. This was the first time Hector had built this plane in a robust way. Previously the construction that had been made was soft.

He first tried to construct several perpendicular bisectors of segment AB by using the corresponding tool or with a perpendicular line to AB through its midpoint D, but it was not possible to follow this procedure in GeoGebra 3D. For this reason, he decided to build two spheres with centers in A and B and radius AB (Figure 2a). Then he created

three points at the intersection of those spheres, as well as the plane determined by these points (figure 2b). In the conversation with the teacher, Hector tried to validate his construction of this plane arguing that *it is a bisector plane... because this plane is perpendicular bisector of AB, so it is perpendicular to AB.*

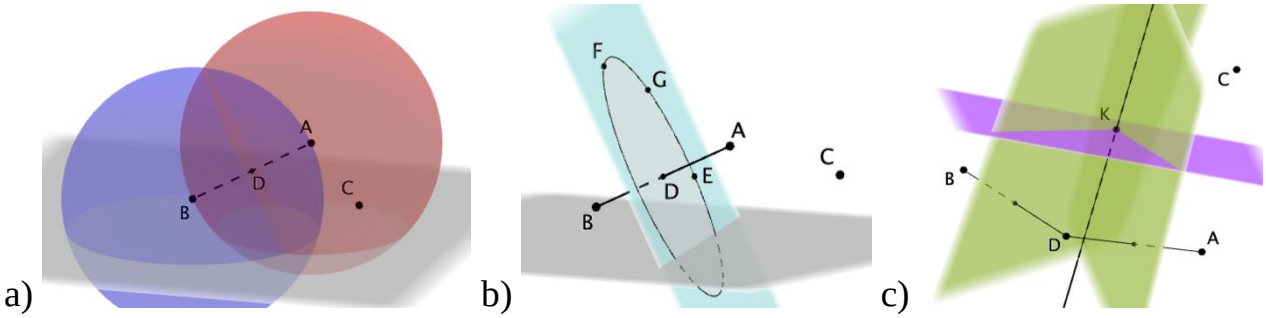


Figure 2. Construction of the bisector plane

Problem 17: Constructing a point equidistant from four non-coplanar points

This problem gave four non-coplanar points A, B, C, and D and asked to construct a point equidistant from all of them. Hector built the bisector planes of points A, D and B, D (green planes, Figure 2c), by using the scheme he created when solved problem 10, and the line intersection of these planes. He then built the bisector plane of points A, C in the same way (purple plane, Figure 2c), determined the point K of intersection between the last bisector plane and the line previously created. Hector stated that point K was the solution to the problem.

To prove the correctness of the construction, Hector assured that any point in the intersection line of the green bisector planes is equidistant from A, B, and D: *I knew that, on this line, these three points* [A, B, D] *would be at the same distance from K ... Because it was the intersection of the bisector planes of A, D and D, B*). Then, he mentioned that his reason to build the third (purple) bisector plane was to determine a set of points equidistant from A and C: ... *I have made the perpendicular bisector* [bisector plane] *so that C and A were* [at] *equal* [distance from any point in the bisector plane]. Hector stated that point K was also equidistant from C, since *If C* [distance CK] *is equal to A* [distance AK], *A is equal to B* [distance BK] *and A is equal to D* [distance DK], *then C is equal to B, C is equal to D…*

ANALYSIS AND DISCUSSION

We have presented a trajectory that began with the procedure that Hector elaborated to build the perpendicular bisector of a 2D segment. Solving problem 07, he recreated on GeoGebra a well-known construction with ruler and compass with the help of the Circumference (Center-Radius) artifact [Sas1]. This procedure was also used for the construction of an equilateral triangle, which led Hector to discover new possibilities for this artifact [Tsa1]. Although this procedure came from his school experience and he did not know why it was valid, the conversation with the teacher and his geometric knowledge allowed him to elaborate a proof of the construction of the perpendicular bisector with circles, as well as that of the equilateral triangle.

When Hector moved to GeoGebra 3D and tried to build an equilateral triangle given one of its sides, he discovered that the scheme that had been useful in 2D was now not available, due to the limitations of the tools in GeoGebra 3D [Tsa2]. To overcome this difficulty, Hector modified the 2D scheme by replacing the circles by spheres, and he obtained a way to construct perpendicular bisectors in 3D and the requested triangle [Sas3]. Hector justified that that construction was correct, using an explicit correspondence between circles and spheres properties, as support for his deductions.

In problem 10, Hector had to modify the scheme again when he built the bisector plane because he did not have the tools used in 2D [Sas3]. This construction, however, was based on the scheme that Hector had elaborated to obtain a perpendicular bisector in 3D and the characterization of this plane as a set of all perpendicular bisectors of the segment. The latter is evidenced in the argumentation he made for the validity of the construction. Hector was modifying the utilization scheme to adapt it to the requirements of each new problem [Tsa3].

Hector's solution to last problem (17) showed a joint use of the modified schemes and the properties of the geometric objects represented through them: on the one hand, the geometric properties of the constructed planes allowed him to guarantee the equidistance of a set of points with respect to three other points; on the other hand, it is the equidistance provided by this scheme what mobilizes the construction of another bisector plane with which the problem was solved.

Hector's solutions to the problems presented, and others that we cannot mention due to the limited length of this paper, show the uses he made of utilization schemes to build bisector planes and perpendicular bisectors, according to the proposed problems [Sas2]. The recurrent use of those schemes to solve different problems [Sas4] and the confidence with which Hector referred to the results obtained through the schemes [Tsa4], revealed in a global way the relationship between him and the circle-sphere artifact. It provides elements to ensure the emergence of an instrument.

The case that we have analyzed provides evidence showing that solving construction-and-proof problems in a 3D-DGE induced the development of proving skills in this mathematically gifted student (we have also obtained similar results with the other students participating in the research experiment). We have showed some glimpses of Hector's advance in his instrumental activity with GeoGebra 3D and the ways it supported the development of his proving skills, since it allowed him to evoke mathematical meanings of represented objects on screen to prove the validity of his constructions.

The nature of this study does not allow generalize its results, but it opens a direction of research to better understand the differentiating characteristics of the processes of learning to prove by mathematically gifted students. A natural continuation of this study is to make similar experiments with average students of the same ages or grades as the mathematically gifted students participating in this experiment. On the other

side, it is necessary to experiment with other 3D geometric relations as well, to provide a broader view of the influence of 3D-DGE on the learning of proof.

ACKNOWLEDGMENT

This paper is part of the R+D+I project PID2020-117395RB-I00, and the predoctoral grant EDU2017-84377-R both funded by MCIN/AEI/10.13039/501100011033.

REFERENCES

Drijvers, P., Kieran, C., Mariotti, M. A., Ainley, J., Andresen, M., Chan, Y. C., Dana-Picard, T., Gueudet, G., Kidron, I., Leung, A., & Meagher, M. (2009). Integrating technology into mathematics education: theoretical perspectives. In C. Hoyles & J.-B. Lagrange (Eds.), *Mathematics education and technology - Rethinking the terrain* (pp. 89-132). Springer.

Fiallo, J., & Gutiérrez, A. (2017). Analysis of the cognitive unity or rupture between conjecture and proof when learning to prove on a grade 10 trigonometry course. *Educational Studies in Mathematics*, *96*(2), 145–167.

Gutiérrez, A., & Jaime, A. (2015). Análisis del aprendizaje de geometría espacial en un entorno de geometría dinámica 3-dimensional. *PNA*, *9*(2), 53-83.

Mariotti, M. A. (2012). Proof and proving in the classroom: dynamic geometry systems as tools of semiotic mediation. *Research in Mathematics Education*, *14*(2), 163-185.

Mariotti, M. A. (2019). The contribution of information and communication technology to the teaching of proof. In G. Hanna, D. A. Reid, & M. de Villiers (Eds.), *Proof technology in mathematics research and teaching* (pp. 173-195). Springer.

Rabardel, P. (1995). *Les hommes et les technologies. Une approche cognitive des instruments contemporains*. Armand Colin.

Sinclair, N., & Robutti, O. (2013) Technology and the role of proof: the case of dynamic geometry. In M. A. Clements, A. J. Bishop, C. Keitel, J. Kilpatrick, & F. K. S. Leung (Eds.), *Third international handbook of mathematics education* (pp. 571-596). Springer.

Sinclair, N., Bartolini Bussi, M. G., de Villiers, M., Jones, K., Kortenkamp, U., Leung, A., & Owens, K. (2016). Recent research on geometry education: an ICME-13 survey team report. *ZDM Mathematics Education*, *48*(5), 691–719.

Sua, C. & Camargo, L. (2019). Geometría dinámica y razonamiento científico: dúo para resolver problemas. *Educación Matemática*, *31*(1), 7-37.

Vergnaud, G. (1996). Au fond de l'apprentissage, la conceptualisation. In R. Noirfalise, & M.-J. Perrin (Eds.), *Actes de l'École d'Été de Didactique des Mathématiques* (pp. 174-185). IREM, Université de Clermont-Ferrand II.

CHINESE LANGUAGE LEARNERS' READING COMPREHENSION WHEN SOLVING MATHEMATICAL WORD PROBLEMS

Emily S. W. Sum[1], Miranda, K. Y. Wong[2], Antonia, Y. T. Yip and Wee Tiong Seah[1]

[1]Melbourne Graduate School of Education, The University of Melbourne, Australia
[2]Department of Psychology, University of Warwick, UK

This exploratory study has been designed to investigate Chinese language learners' (CLLs) reading comprehension when solving additive word problems. 200 primary one students (including 59 CLLs) in eight Hong Kong mainstream schools were asked to translate five Chinese word problems into diagrams or arithmetic forms (number sentences or equations). Results show that local Cantonese-speaking students generally outperformed CLLs in translating the word problems correctly to a diagram and/or mathematical expression. We argue that there is a need for early intervention to support CLLs to acquire reading comprehension skills in mathematics learning.

INTRODUCTION

In Hong Kong, most ethnic minority (mainly South-East Asian) students attend mainstream schools, which largely use Chinese as the language of instruction (that is, the Cantonese dialect as the spoken form and traditional Chinese characters as the written code). These linguistic minority students speak different home languages other than Chinese, and we refer to them as Chinese language learners (CLLs) in this study. Most teachers were not trained in Second Language (L2) pedagogies (Tsung et al., 2010), and they mainly use traditional methods of teaching Chinese as first language (L1) with "one size fits all" pedagogy. CLLs have frequently reported to encounter great difficulties in acquiring Chinese language (Ku et al., 2006; Shum et al., 2016), and learning mathematics in a second/additional language (Sum et al., 2022; Tse & Hui, 2012). They are learning mathematics in Chinese while simultaneously learning Chinese as L2 at school.

Moreover, in mathematics classrooms, the colloquial Cantonese used in teaching and the Chinese characters in textbooks are sometimes different. The choice of lexical items reflects the level of formality of utterances, for instance, 食咗/吃了 (*ate*), 剩返/還剩下 (*left*). "[T]he colloquial form that they grow up speaking and occasionally see being written but are not explicitly taught to speak, read, or write" are different from "the formal, standard variety that they are taught to read and write at school" (Bauer, 2016, p. 121). This creates problems for CLLs and contributes to the complicated language situation in schools. Chinese is an ideogram, which is substantially different from alphabetical orthographies used in English and some South-East Asian languages. In this context, this study investigates students' reading comprehension, and analyses the language aspects of solving mathematical word problems.

2023. 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. 251-258). PME 46.

LITERATURE REVIEW

In order to succeed in solving word problems, students require more than arithmetic computational skills, and reading comprehension has long been found as a critical factor (Cummins et al., 1988; Fuchs et al., 2018; Vilenius-Tuohimaa, 2008). Previous research also suggests that working with word problems presents additional challenges for language learners (Barwell, 2005), including language complexity (vocabulary, sentence structure) (Abedi et al., 2000); and unfamiliar background or sociocultural knowledge of the word problems (Martiniello & Wolf, 2012).

Mathematical word problem is a linguistic genre, characterised by having a pragmatic structure – a set-up, an information, and a question (Gerofsky, 1996). Morgan (2006) used Halliday's Systemic Functional Linguistic (SFL) to analyse the nature of mathematical language. The ideational function of the text within the transitivity system is "the types of processes, the participants in those processes and the representation of actors" (Morgan, 2006, p. 227) and the notion of mathematics register is "the meanings that belong to the language of mathematics" (Halliday, 1978, p. 195).

Research shows that Hong Kong students had difficulties in writing number sentences when solving word problems (Wong & Ho, 2017). The language demand towards comprehending the word problems is high, even for native Cantonese speakers (Ng et al., 2021). At primary one, students would encounter lexical items identifiable as mathematical. For instances, 加法 (*addition*), 加號 (*addition sign*), 符號 (*symbol*), 等號 (*equal sign*). Each character in these words represents a distinct morpheme (加 means add, 等 means equal, 號 means symbol). The words 加號 and 等號 share the same orthographic form of 號, but their meanings are inferred from their respective component morphemes (addition sign / equal sign). Students need to understand the morphological structure within the compound words, which is unique to the Chinese language (Mathews & Yip, 2011). Also, Cantonese is a numeral classifier language, in which modification of count nouns with numerals always requires a classifier. When solving word problems, classifiers, categorical in nature, provides students with a meaning for the numbers and a better sense of different numbers in the problem situation (Sum & Kwon, 2018).

While there is an increased attention to the issues related to the teaching and learning of mathematics for English language learners over the past 30 years (de Araujo et al., 2018), little has been done in relation to the Chinese language, and even less empirically explored in the Cantonese-speaking context. CLLs' reading comprehension of word problems remains unexamined in the current literature. Considering that, as a field of study in language and communication in mathematics education, the case of Cantonese Chinese as L2 is worthy of attention. We thus pose the following research questions in this study:

RQ1: What is the performance gap of local Cantonese-speaking students and CLLs' reading comprehension of word problems, if any?

RQ2: What are the linguistic features of word problems in Chinese language that may affect the way meaning is constructed by Cantonese-speaking students and CLLs?

RESEARCH METHODS

Participants

Participants were 200 primary one students (59 of whom were CLLs) from eight schools in Hong Kong, with Cantonese as the language of instruction. This grade level is selected because policy document indicates that "ethnic minority students were very much weaker in Chinese and slightly weaker in Mathematics than their Chinese counterparts at the point of Primary 1 admission" (Kapai, 2015, p. 27). Therefore, primary one is a good starting point to examine the gap, if any, between local students and CLLs.

Data collection and analysis

The research began in June 2022 with the collection of data related to students' initial reading comprehension of additive word problems. Students were asked to read and interpret the sentences mathematically, translating the problem text into drawings or diagrams, and/or writing down the associated number sentence or equation. The written context consisted of five additive word problems, which are classified based on the semantic structure of the arithmetic word problems (Riley et al., 1983). The test items were derived from a primary one mathematics textbook. They included one Combine question, one 1-step and one 2-step Change questions, and two Compare questions. The number sentences of the word problems were in standard form, i.e., $A + B = (\quad)$ and $A - B = (\quad)$. All word problems were written in traditional Chinese characters. We asked the students to indicate which word(s) they did not understand by underlining the words/phrases, and then to draw a picture/diagram or write down a number sentence/equation to represent the mathematical content of the word problems. Students were not required to calculate the numerical answers to the problems as it is beyond mathematising the problem text.

Each word problem was analysed using traditional grammar (lexical, syntactic, and semantic features) and the SFL framework (Halliday, 2014) – the ideational function and the grammatical resources that realise its meanings – to illustrate the language features of Cantonese language (Shum & Mickan, 2019). Our focus was on how language is integrated with content, analysing the linguistic challenges CLLs encounter in comprehending the text in the word problems, and extracting the numbers and relating them through arithmetic operations.

Descriptive statistics of, and comparison (chi-square tests) between, the key measures of the study were calculated. Two repeated measures ANOVAs were conducted to examine the correct response rate across the five mathematical word problems for local students and CLLs, respectively, with Question as the within-subject independent

variable and Correct Response Rate as dependent variable. To further investigate students' comprehension of the word problems using different methods, a series of 2×2 repeated measures ANOVA was conducted for each question to examine the correct response rate of using different methods to represent the problem text between local students and CLLs for each item, with Method (diagram, equation) as the within-subject independent variable, Home Language (Chinese, Non-Chinese) as the between-subject independent variable, and Accuracy Rate as dependent variable.

RESULTS AND DISCUSSION

Students' performance

Table 1 shows the correct response rate of each word problem. There were large differences in correct response rates between different types of problems. In solving Combine problem, local students performed significantly better than CLLs, χ^2 (1) = 14.47, p < .001, either by drawing a correct diagram, writing a correct number sentence/equation, or working out a correct answer. Yet, no significant differences between local students and CLLs were found in Change and Compare problems.

	Correct Response Rate (%)			
Question Type	Local (n=141)	CLLs (n=59)	χ^2	p
1 Combine (value unknown)	86.5%	62.7%	14.47	< .001
2 Change (result unknown)	72.3%	62.7%	1.82	.177
3 Change (result unknown)	57.4%	47.5%	1.67	.196
4 Compare (difference unknown)	37.6%	33.9%	.244	.621
5 Compare (compared quantity unknown)	52.5%	52.5%	.00	.994

Table 1: Percentage of correct response between local students and CLLs

Across the five word problems, local students and CLLs showed similar patterns in their correct response rates, with highest correct response rate for Q1 and lowest correct response rate for Q4. The results of repeated measures ANOVAs revealed significant effect for Question for local students, $F(4, 137) = 38.90$, $p < .001$, $\eta_p^2 = .53$, and for CLLs, $F(4, 55) = 4.56$, $p = .003$, $\eta_p^2 = .25$. For local students, pairwise comparisons showed that the correct response rate of Q1 was higher than that of the other four questions ($ps < .01$), while the correct response rate of Q4 was lower than that of the other four questions ($ps < .05$). For the remaining three questions, the correct response rate of Q2 was higher than that of Q3 and Q5 ($ps < .05$), while the correct response rates of Q3 and Q5 were comparable ($p = 1.000$). For CLLs, pairwise comparisons showed that the correct response rate of Q4 was lower than that of Q1 and Q2 ($ps < .05$), while the correct response rates of other questions were comparable ($ps > .05$).

A series of 2×2 repeated measures ANOVA was conducted for each question to further investigate local students and CLLs' comprehension of mathematical word problems using different methods. Table 2 summarises the results.

		F	*p*	η_p^2
Q1	Method	$F(1, 198) = 114.05$	< .001	.37
	Home Language	$F(1, 198) = 2.24$	.136	.01
	Method × Home Language	$F(1, 198) = 6.04$	.015	.03
Q2	Method	$F(1, 198) = 93.01$	< .001	.32
	Home Language	$F(1, 198) = .58$	.449	.003
	Method × Home Language	$F(1, 198) = 4.83$	.029	.02
Q3	Method	$F(1, 198) = 72.56$	< .001	.27
	Home Language	$F(1, 198) = 1.83$	.178	.01
	Method × Home Language	$F(1, 198) = 4.34$	.038	.02
Q4	Method	$F(1, 198) = 72.32$	< .001	.27
	Home Language	$F(1, 198) = .09$	.770	< .001
	Method × Home Language	$F(1, 198) = .05$	.817	< .001
Q5	Method	$F(1, 198) = 72.65$	< .001	.27
	Home Language	$F(1, 198) = .21$	.648	.001
	Method × Home Language	$F(1, 198) = .79$	.374	.004

Table 2: Summary of repeated measures ANOVAs for comparisons of correct response rate of local students and CLLs

As shown in Table 2, for all items, the main effect for Method was significant (*ps* < .001), while the main effect for Home Language was non-significant (*ps* > .05). Specifically, for each question, students applying equation to represent the text had higher accuracy rate than students applying diagram to represent the text. Significant interaction effect of Method × Home Language was only found for Q1, Q2, and Q3 (*ps* < .05). The interaction patterns were slightly different across the three questions and the pairwise comparisons will be discussed in the following section.

The ideational function and grammar of the word problems

For Q1 Combine problem, local students (Mean = .70, SD = .46) had significantly higher accuracy rate than CLLs (Mean = .53, SD = .50) when using equation, $p = .027$. In this problem, one clause used relational process to describe the initial quantity, with an additive conjunctive cohesion 和 (*and*), followed with a question of relational process asking the combined value. 5.67% of local students and 11.86% of CLLs

indicated that they did not know the mathematical vocabulary 共有(*together*) but made good use of the morphological cue, 和(*and*), in the sentence to decode the problem.

For Q2 one-step Change problem, local students (Mean = .59, SD = .49) and CLLs (Mean = .46, SD = .50) who used equation had comparable accuracy rate, p = .142. However, for Q3, two-step Change problem, local students (Mean = .48, SD = .50) had marginally significant higher accuracy rate than CLLs (Mean = .34, SD = .48) when using equation, p = .072. In these problems, there were two to three clauses starting with either relational or existential process of the initial quantity, followed by a material process to describe the change of quantity, and ended in a question of existential process to find the resulting quantity. The words indicating the change in quantities, such as 吃了(*ate*), 還剩下(*left*), 多買(*bought*), and 再買(*bought again*), were not colloquial Cantonese, and some students have difficulties recognising these characters to comprehend the meaning conveyed in the text, and to mathematise the additive situation of the problems. 15.60% of local students and 18.64% of CLLs indicated that they did not understand the phrase 還剩下(*left*). Also, students found it more difficult when there were multiple ways of saying the same operation, including but not limited to 和 共有 and 現有, all of which can refer to finding the sum, especially when they share the same morphemes (有 in 共有, 現有).

In Q4 and Q5 Compare problems, there were two to three clauses in either relational or existential process to show the initial quantity, followed by a comparison circumstance in material process to illustrate the differences between the qualities, and ended with a question of relational process. Significant effect was found for Method in these Compare problems, in which students who used equation had higher accuracy rates than students who used diagram. The accuracy rates for both questions were similar for both local students and CLLs. This may be due to "the knowledge needed specially for compare problems develops later than the specific knowledge involved in the combine and change problem" (Riley & Greeno, 1988, p.71). Another possible explanation is that the comparative adverbs such as 相差(*difference*) and 比… 多(*more than*) used were not easy to understand. Specifically, 9.93% of local students and 16.95% of CLLs expressed that they did not know the phrase 相差(*difference*). The language complexity of Compare problems is not only at the word level, but sentence level (syntactic structures) as well, given that word order is crucial for interpretation in Chinese (Chang, 1992). It is worth noting that students who did not understand the comparative adverb 相差 (*difference*) in the material process might have miscomprehended it for the presence of the conjunction 和(*and*) and therefore chose the wrong operation.

CONCLUSION

To conclude, Hong Kong local students' reading comprehension in solving mathematical word problems is generally better than CLLs'. When compared to Cantonese-speakers, more CLLs have difficulties in understanding the ideational functions of the problem text in order to mathematise the additive situations. Early

intervention to support CLLs to acquire reading comprehension skills in mathematics learning is thus needed. Despite having no information on students' strategies in translating the words into arithmetic forms, the findings of the present study provide important directions for further research to support intervention approaches, which are to identify the characteristics of Cantonese that impact on CLLs' meaning-making in mathematics learning, and to explore how the linguistic/lexical complexity of mathematical word problems impacts on the way CLLs process written text. This should allow us to maintain the rigor of mathematical problems while providing equal access to learning for all students.

REFERENCES

Abedi, J., Lord, C., Hofstetter, C., & Baker, E. (2000). Impact of accommodation strategies on English language learners' test performance. *Educational Measurement*, *19*(3), 16-26.

Bauer, R. S. (2016). The Hong Kong Cantonese language: Current features and prospects. *Global Chinese*, *2*(2), 115-161. https://doi.org/10.1515/glochi-2016-0007

Barwell, R. (2005). Working on arithmetic word problems when English is an additional language. *British Education Research Journal*, *31*(3), 329-348.

Chang, H.-W. (1992). The acquisition of Chinese syntax. In H.-C. Chen & O. J. L. Tzeng (Eds.), *Language processing in Chinese* (pp. 277-311). North-Holland.

Cummins, D. D., Kintsch, W., Reusser, K., & Weimer, R. (1988). The role of understanding in solving word problems. *Cognitive Psychology*, *20*(4), 405-438. https://doi.org/10.1016/0010-0285(88)90011-4

de Araujo, Z., Roberts, S. A., Willey, C., & Zahner, W. (2018). English learners in K-12 mathematics education: A review of the literature. *Review of Educational Research*, *88*(6), 879-919.

Fuchs, L. S., Gilbert, J. K., Fuchs, D., Seethaler, P. M., & Martin, B. N. (2018). Text comprehension and oral language as predictors of word-problem solving: Insights into word-problem solving as a form of text comprehension. *Scientific Studies of Reading*, *22*(2), 152-166. https://doi.org/10.1080/10888438.2017.1398259

Gerofsky, S. (1996). A linguistic and narrative view of word problems in mathematics education. *For the Learning of Mathematics*, *16*(2), 36-45.

Halliday, M. A. K. (1978). *Language as social semiotic: The social interpretation of language and meaning.* Edward Arnold.

Halliday, M. A. K. (2014). *Halliday's introduction to functional grammar* (4th ed.). Routledge.

Kapai, P. (2015). The education of ethnic minorities. In P. Kapai (Ed.), *The status of ethnic minorities in Hong Kong 1997- 2014* (pp. 124-175). Centre for Comparative and Public Law, Faculty of Law, The University of Hong Kong.

Ku, H. B., Chan, K. W., & Sandhu, K. K. (2006). *A research report on the employment of South Asian ethnic minority groups in Hong Kong*. Centre for Social Policy Studies, Department of Applied Social Sciences, The Hong Kong Polytechnic University.

Martiniello, M., & Wolf, M. K. (2012). Exploring ELLs' understanding of word problems in mathematics assessments: The role of text complexity and student background knowledge. In S. Celedón-Pattichis & N. Ramirez (Eds.), *Beyond good teaching: Strategies that are imperative for English language learners in the mathematics classroom* (pp. 151-162). NCTM.

Matthews, S., & Yip, V. (2011). *Cantonese: A comprehensive grammar* (2nd ed.). Routledge.

Morgan, C. (2006). What does social semiotics have to offer mathematics education research? *Educational Studies in Mathematics, 61*(1/2), 219-245.

Ng, W.-S., Wong, T. T.-Y., & Fong, C. Y.-C. (2021). Contributions of reading comprehension subsills to arithmetic word-problem solving among Chinese primary school students. *Journal of Cognition and Development, 22*(4), 585-604.

Riley, M. S., & Greeno, J. G. (1988). Developmental analysis of understanding language about quantities and of solving problems. *Cognition and Instruction, 5*(1), 49-101. https://doi.org/10.1207/s1532690xci0501_2

Riley, M. S., Greeno, J. G., & Heller, J. I. (1983). Development of children's problem-solving ability in arithmetic. In H. Ginsberg (Ed.), *The development of mathematical thinking* (pp. 153-196). Academic Press.

Shum, M. S. K., Gao, F., & Ki, W. W. (2016). School desegregation in Hong Kong: Non-Chinese linguistic minority students' challenges to learning Chinese in mainstream schools. *Asia Pacific Journal of Education, 36*(4), 533-544.

Shum, M. S. K., & Mickan, P. (2019). *Researching Chinese language education: Functional linguistic perspectives*. Routledge.

Sum, E. S. W., & Kwon, O. N. (2018). An analysis of linguistic features of multiplication tables and the language of multiplication. *EURASIA Journal of Mathematics, Science and Technology, 14*(7), 2839-2856. https://doi.org/10.29333/ejmste/90760

Sum, E. S. W., Wong, K. L., & Seah, W. T. (2022). Attributes of mathematics learning valued by ethnic minority students in Hong Kong mainstream schools. *Asia Pacific Journal of Education*, 1-15. https://doi.org/10.1080/02188791.2022.2071215

Tse, S. K., & Hui, S. Y. (2012). Supporting ethnic minority students learning the Chinese language in multilingual Hong Kong. *L1-Educational Studies in Language and Literature, 12*(2), 1-27. https://doi.org/10.17239/L1ESLL-2012.02.09

Tsung, L., Zhang, Q., & Cruickshank, K. (2010). Access to majority language and educational outcomes: South Asian background students in postcolonial Hong Kong. *Diaspora, Indigenous, and Minority Education, 4*(1), 17-32.

Vilenius-Tuohimaa, P. M., Aunola, K., & Nurmi, J.-E. (2008). The association between mathematical word problems and reading comprehension. *Educational Psychology, 28*(4), 409-426.

Wong, T. T.-Y., & Ho, C. S.-H. (2017). Component processes in arithmetic word-problem solving and their correlates. *Journal of Education Psychology, 109*(4), 520-531. https://doi.org/10.1037/edu0000149

ENHANCING SPATIAL REASONING THROUGH GEOMETRY TRANSFORMATION INSTRUCTION IN GHANA

Mawuli Kofi Tay, Armando Paulino Preciado Babb

University of Calgary, Calgary

Spatial reasoning has been identified as a key factor in learning mathematics; yet, efforts to enhance students' spatial reasoning at the high school level are still scarce. This paper reports the results of a study intended to compare the effects of dynamic versus static visualization instruction on high school students' spatial reasoning skills(n=77) during a 4-week intervention. While the analysis showed no significant differences between the groups, both instructional approaches had a significant impact on the three components of students' spatial reasoning skills (mental rotation, spatial orientation, and spatial visualization), suggesting that students' spatial reasoning can be enhanced within the context of geometry transformations when an appropriate pedagogical approach informs instruction.

INTRODUCTION

The term spatial reasoning has multiple definitions depending on research disciplinary perspectives and purposes. However, there is an overarching consensus that spatial reasoning involves the ability to mentally manipulate or transform objects being visualized (Bruce et al., 2017; Davis & Francis, 2020). It is well documented in both educational psychology and mathematics education literature that spatial reasoning improves mathematics achievement (Bruce et al., 2017; Davis et al., 2015; Mulligan et al., 2018). Collectively, there is consensus that spatial reasoning can be developed through training or targeted instruction (Uttal et al., 2013). In doing so, the instruction should be tailored towards specific learning objectives. Yet, scholars have paid little attention to specific instruction within the mathematics classroom that purposefully develops and improves students' spatial reasoning despite the various calls from notable mathematical associations and scholars to integrate spatial reasoning activities into K-12 mathematics instructions (Davis et al., 2015; Patahuddin et al., 2020). Even though there are comparatively few scholars who are reluctantly trying to incorporate spatial reasoning in mathematics classrooms (see Adams et al., 2022; Hawes et al., 2017), spatial reasoning remains underrepresented in many mathematics instructions, as it is the case of education in Ghana. Accordingly, this study seeks to bridge this gap comparing the effect of different approaches to visualization—static vs dynamic— on students' spatial reasoning skills in a Ghanaian high school.

Spatial reasoning and geometric transformations are intrinsically related in terms of mental manipulation or transformation of objects. For this reason, this study assumes that an instructional unit of geometry transformations has the potential to improve spatial reasoning. Arguably, by demonstrating the motions involved in geometric transformations (i.e., rotation, reflection, and translation), the teacher might be able to provide students with visualization skills for imagining geometric transformations.

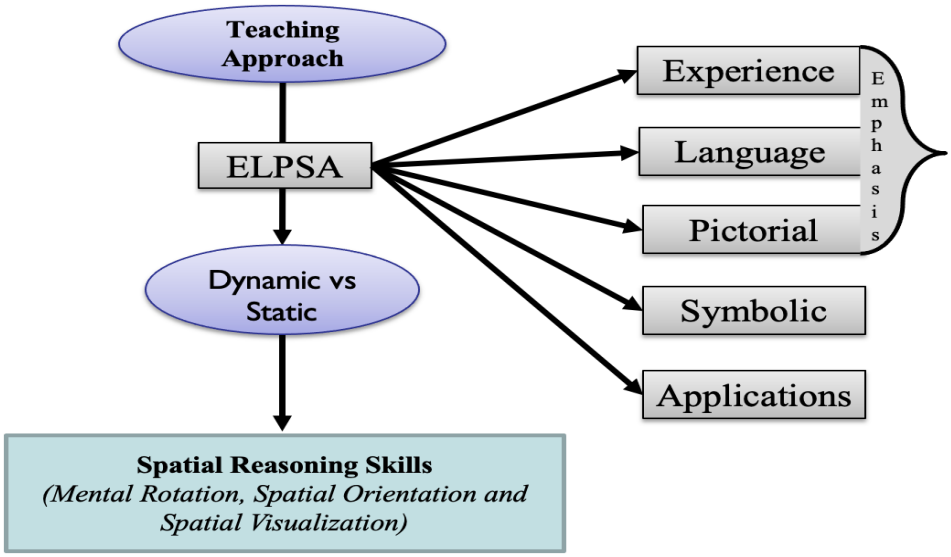
2023. In M. Ayalon, B. Koichu, R. Leikin, L. Rubel & M. Tabach (Eds.). *Proceedings of the 46th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 259-266). PME 46.

Earlier researchers have indicated that spatial reasoning is not a single unitary construct, but it comprises multiple constructs (Linn & Petersen, 1985; Lohman, 1979). In a more recent study, Ramful et al. (2017) categorized spatial reasoning into a three-tier framework comprising: mental rotation, spatial orientation, and spatial visualization. It is worth noticing from these descriptions that spatial visualization, spatial orientation, and mental rotation are common among multiple compositions of spatial reasoning. For this reason, this study proposes to use Ramful and colleagues' three-tier framework test of spatial reasoning. That will allow us to measure several dimensions of spatial reasoning of students in a single test rather than implementing separate tests on spatial visualization, mental rotation, and spatial orientation.

CONCEPTUAL FRAMEWORK

Figure 1 summarizes both the relationship between variables in the study—type visualization as the independent variable and spatial reasoning skills as dependent variable—and the pedagogical framework guiding the teaching approach for both groups: dynamic and static visualizations. The design for the lessons in both cases was based on the Experience-Language-Pictorial-Symbolic-Application (ELPSA) framework (Lowrie & Patahuddin, 2015). The *experience* draws on what the students know about the topic in relation to their own experiences within and outside the classroom contexts. The *language* focuses on how appropriate terminology is used to represent mathematical ideas. The *pictorial* outlines learning around the use of visual representations to represent mathematical ideas. The *symbolic* is aligned with the use of symbols to formalize ideas or concepts. The *applications* focus on how the knowledge obtained can be applied in different situations. This study emphasizes the experience, the language, and the pictorial.

Figure 1: Conceptual Framework



METHODOLOGY

Research Design

A quasi-experimental design with pre-test/post-test control group was employed in the study. The study was carried out over a period of 4 weeks using the dynamic visualization instructional approach in the experimental group and the static visualization instructional approach in the control group. While the independent variable was the mode of visualization, the dependent variable was the students' test scores on spatial reasoning skills assessed using Ramful and colleagues' (2017) Spatial Reasoning Instrument before and after the 4 weeks period.

Participants

Participants for this study were students (16 – 20 years old) from two classes at Delana Senior High School (DSHS) (pseudonym) located at Ho in the Volta Region of Ghana. The researcher worked with a sample of 77 students who gave their consents and participated in both tests. This number consists of 42 students in the class A and 35 students in class B who were purposively assigned at the classroom level to either the experimental (n = 35) or control (n = 42) group.

Procedure

The teaching approach for both experimental and control groups was the same, except for the use of dynamic and static representations. The dynamic and static visualization instruction to geometry transformation refer to the teaching approaches involving teacher-led demonstrations using GeoGebra as a visual instructional material. In dynamic visualization instruction, animation and dragging were applied as compared to static visualization instruction where no animation or dragging were applied. The lessons of the unit were codesigned with the teacher following the ELPSA teaching and learning framework. Lessons for both groups were held in the students' regular mathematics classroom using projector and screen. In both instructions, students were taught concepts of geometry transformation such as *rotation, reflection, and translation*. Students' hands-on activities were designed using lived experiences about geometry, worksheets hand in hand with the lesson plan by the researcher and teacher. This instructional approach for both groups also included collaborative learning.

Instruments

Ramful and colleagues' (2017) Spatial Reasoning Instrument consists of 10 multiple-choice items for each of the three constructs, totalling 30 points for the overall score. The *spatial visualization* component involves symmetry, patterns, 2D and 3D shapes and their relationships, part-whole relationships, reflection, and symmetry; the *spatial orientation* component involves orienting oneself in space, readings, the relationship between cardinal points and viewing images or objects from the front, top or side; and the *mental rotation* component involves rotation of 2D and 3D objects, clockwise and anticlockwise directions. The instrument was adapted following a pilot study within the Ghanaian context.

Data Analysis

Prior to data analysis, erroneous entries and missing values were checked, and data were cleaned. The following assumptions (Tabachnick & Fidell, 2019) were verified: normality, homogeneity of variance, interval data, independence, homogeneity of the regression of slopes, and no outliers. Based on the results of verifying these assumptions, a paired sample t-test was used for the pre-test and post-test comparisons within the groups and an analysis of covariance (ANCOVA) was used to determine differences between the dynamic and static groups. The statistical analysis was conducted using IBM SPSS version 28.

RESULTS

The paired sample *t*-test results, in Table 1, reveal a statistically significant increase in students' performance from the pre-tests to the post-tests within both dynamic and static groups across the various spatial reasoning components—mental rotation, spatial orientation, and spatial visualization—indicating a large effect size. These results implied that after students had gone through the treatment, they improved significantly in all three components of spatial reasoning skills.

Table 1: Paired Sample t-Test Results of Pre-Test and Post-Test Scores by Group

Scores	Groups	Tests	Mean	SD	*t*	*df*	*p*	Cohen's d
Mental Rotation	Dynamic	Pre	4.37	1.40	8.99	34	<.001	1.52
		Post	7.23	1.66				
	Static	Pre	4.60	1.88	10.29	41	<.001	1.59
		Post	7.40	1.50				
Spatial Orientation	Dynamic	Pre	6.51	1.44	7.32	34	<.001	1.24
		Post	8.46	1.17				
	Static	Pre	6.31	1.68	5.17	41	<.001	0.80
		Post	7.86	1.65				
Spatial Visualization	Dynamic	Pre	3.40	1.67	7.98	34	<.001	1.35
		Post	6.34	1.71				
	Static	Pre	3.38	1.82	9.07	41	<.001	1.40
		Post	6.02	1.42				
Spatial Reasoning (Combined Scores)	Dynamic	Pre	14.29	2.85	15.30	34	<.001	2.59
		Post	22.03	3.02				
	Static	Pre	14.36	3.49	12.85	41	<.001	1.98
		Post	21.29	3.45				

The ANCOVA results from Table 2 show that there was no statistically significant mean difference in the gain scores of students in the dynamic and static group with respect to the mental rotation [F(1, 74) = 0.08, p = 0.774], spatial orientation [F(1, 74) = 2.89, p = 0.093], and spatial visualization [F (1, 74) = 0.90, p = 0.345] where the pre-tests were used as a covariate. In total, there was no significant mean difference in the gain scores of students in the dynamic and static group with respect to the overall spatial reasoning test [F (1, 74) = 1.41, p = 0.239] where the pre-test scores of the overall test were used as a covariate.

Table 2: ANCOVA of Spatial Reasoning by Group

Spatial Reasoning	Source	Sum of Squares	df	Mean Square	F	Sig.
Mental Rotation	Pre-test	26.43	1	26.43	12.24	<.001
	Group	0.18	1	0.18	0.08	0.774
	Error	159.86	74	2.16		
	Total	4318.00	77			
Spatial Orientation	Pre-test	14.91	1	14.91	7.72	0.007
	Group	5.58	1	5.58	2.89	0.093
	Error	142.92	74	1.93		
	Total	5254.00	77			
Spatial Visualization	Pre-test	11.42	1	11.42	4.93	0.030
	Group	2.09	1	2.09	0.90	0.345
	Error	171.45	74	2.32		
	Total	3115.00	77			
Spatial Reasoning (Combined Scores)	Pre-test	189.75	1	189.75	23.10	<.001
	Group	11.56	1	11.56	1.41	0.239
	Error	607.79	74	8.21		
	Total	36811.00	77			

DISCUSSION AND CONCLUSION

The present study sought to determine whether there is a difference between students' spatial reasoning test scores through the implementation of in-class activities and lessons using dynamic and static visualization instructional approaches to a unit geometry transformation. The results showed no statistically significant difference between the student's spatial reasoning test scores when taught with dynamic or static visualization instruction. These findings are consistent with previous studies that

showed non-significant difference between dynamic and static visualization instruction (Hegarty, 2004; Tversky et al., 2002). These studies attributed non existing significant difference to transient information nature of dynamic visualization which does not give permanent information as compared to static visualization. Even though the mean difference between dynamic visualization and static visualization groups was not statistically significant, there was statistically a significant improvement in student's spatial reasoning test scores within each group. The baseline results showed that students within each of the dynamic and static groups, demonstrated far-reaching improvements across the three separate measures of spatial reasoning —spatial orientation, mental rotation, and spatial visualization—suggesting that their improvements were not specific to a particular spatial reasoning element. This result concurs with previous studies that found a significant influence of dynamic and static visualization instruction on developing spatial reasoning skills (Baki et al., 2011; Güven & Kosa, 2008). The significant effect on students' spatial reasoning skills might be attributed to the fact that spatial reasoning could be improved through training or targeted instruction (Hawes et al., 2017; Uttal et al., 2013). This possibility is supported by the implementations of geometry transformation lessons that involve identifying rotation, reflection, and symmetry of objects which indirectly demands mental manipulation of spatial information. Engaging students with lived experiences about rotating, translating, reflecting, visualizing changes, and imagining the position of two-dimensional objects after being rotated—which were included in the instructional approach for both groups—could directly be linked to developing students' mental rotation skills.

Another key factor that could influence the significant positive effect of dynamic and static visualization instruction on students' spatial reasoning could be credited to the implementations of the ELPSA pedagogy framework (Lowrie & Patahuddin, 2015) in designing geometry transformation lessons. Several researchers investigating the impact of interventions on transfer between student's spatial reasoning and mathematics achievement do implicitly incorporate a pedagogical framework into planning lessons activities (Mix et al., 2020; Sorby et al., 2013; Xu & LeFevre, 2016). Some of these studies do not find a significant positive transfer between student's spatial reasoning and mathematics achievement (e.g., Xu & LeFevre, 2016). Studies that found a narrow or broader impact involved research participants in various spatial tasks prolonged for more extended periods (Mix et al., 2020; Sorby et al., 2013). However, it is interesting to know that studies that considered a particular pedagogy for the design of lesson activities, such as the present study, had a significant positive transfer on students' spatial reasoning skills (Adams et al., 2022; Hawes et al., 2017; Lowrie et al., 2019). These findings suggest that integrating an appropriate pedagogical framework into designing mathematics classroom interventions could develop students' spatial reasoning skills.

In general, instruction with dynamic and static visualization positively affects students' spatial reasoning skills across all the sub-components. This current study contributes

significantly to the existing literature on improving spatial reasoning among students and methods of training for spatial reasoning. This study's approach to fostering students' spatial reasoning within the context of geometry transformation instruction varies from usual approaches to enhancing students' spatial reasoning skills. Most spatial reasoning training interventions have been comparatively short in duration involving building blocks, paper folding, and cutting, and playing digit games (Uttal et al., 2013). These interventions are primarily designed and implemented by researchers themselves. In contrast, this study was conducted within school mathematics instruction periods for four weeks in duration (16 hours) with varieties of geometry transformation activities that inherently connect to spatial tasks. The lessons were designed in collaboration with the students' regular mathematics teacher and implemented by him. Recognizing these differences helps expound the important descriptions of this study that might have contributed to the observed findings.

The findings from the present study have implications for mathematics educators and curriculum developers: (i) students' spatial reasoning skills can be enhanced in the context of a geometry transformation instruction, (ii) using both dynamic and static visualization instructions with an appropriate teaching and learning framework can enhance students' spatial reasoning skills (iii) mathematics instructional materials could consider teaching and learning activities that encourage students' spatial reasoning skills. However, these implications and generalization cannot be extended beyond the school where the study was conducted, but schools with similar features. Also, the study only focuses on one specific element of geometry and a particular pedagogy (ELPSA). This again indicates that results from this study may not be generalized to other teaching approaches and other topics in geometry. In the further studies, other teaching approaches with difference mathematics content that influence students' spatial reasoning in different grade schools can be explored. Moreover, the relationship between geometry transformation and spatial reasoning is still under investigation.

REFERENCES

Adams, Resnick, I., & Lowrie, T. (2022). Supporting senior high-school students' measurement and geometry performance: Does spatial training transfer to mathematics achievement? *Mathematics Education Research Journal, 1–22*.

Baki, A., Kosa, T., & Guven, B. (2011). A comparative study of the effects of using dynamic geometry software and physical manipulatives on the spatial visualisation skills of pre-service mathematics teachers. *British Journal of Educational Technology, 42*(2), 291-310.

Bruce, C., Davis, B., Sinclair, N., McGarvey, L., Hallowell, D., Drefs, M., Francis, K., Hawes, Z., Moss, J., Mulligan, J., Okamoto, Y., Whitely, W., & Woolcott, G. (2017). Understanding gaps in research networks: using spatial reasoning as a window into the importance of networked educational research. *Educational Studies in Mathematics, 95*, 143–161.

Davis, B., & Francis, K. (2020). "Spatial reasoning" in discourses on learning in education. https://learningdiscourses.com.

Davis, B., & Spatial Reasoning Study Group (Eds.). (2015). *Spatial reasoning in the early years: Principles, assertions, and speculations*. Routledge.

Güven, B., & Kosa, T. (2008). The effect of dynamic geometry software on student mathematics teachers' spatial visualization skills. *Turkish Online Journal of Educational Technology-TOJET, 7(4)*, 100-107.

Hawes, Z., Moss, J., Caswell, B., Naqvi, S., & MacKinnon, S. (2017). Enhancing children's spatial and numerical skills through a dynamic spatial approach to early geometry instruction: effects of a 32-week intervention. *Cognition and Instruction, 35*, 1–29.

Hegarty, M. (2004). Dynamic visualizations and learning: Getting to the difficult questions. *Learning and Instruction, 14*, 343–351.

Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development, 56*, 1479-1498.

Lohman, D. (1979). Spatial ability: Review and re-analysis of the correlational literature (Tech. Rep. No. 8). Stanford University School of Education.

Lowrie, T., & Patahuddin, S. M. (2015). ELPSA as a lesson design framework. *Journal of Mathematics Education, 6(2)*, 1–15.

Mix, K. S., Levine, S. C., Cheng, Y. -L., Stockton, J. D., & Bower, C. (2020). Effects of spatial training on mathematics in first and sixth grade children. *Journal of Educational Psychology, 113*(2), 304–314.

Mulligan, J., Woolcott, G., Mitchelmore, M., & Davis, B. (2018). Connecting mathematics learning through spatial reasoning. *Mathematics Education Research Journal, 30*, 77–87.

Patahuddin, S. M., Rokhmah, S., & Ramful, A. (2020). What does teaching of spatial visualisation skills incur: an exploration through the visualise-predict-check heuristic. *Mathematics Education Research Journal, 32(2)*, 307-329

Ramful, A., Lowrie, T., & Logan, T. (2017). Measurement of spatial ability: Construction and validation of the spatial reasoning instrument for middle school student. *Journal of Psychoeducational Assessment, 35(7)*, 709–727.

Sorby, S., Casey, B., Veurink, N., & Dulaney, A. (2013). The role of spatial training in improving spatial and calculus performance in engineering students. *Learning and Individual Differences, 26*, 20-29.

Tabachnick, B. & Fidell, L. (2019). Using multivariate statistics (7th ed.). Pearson.

Tversky, B., Morrison, J. B., & Betrancourt, M. (2002). Animation: Can it facilitate? *International Journal of Human–Computer Studies, 57*, 247–262.

Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin, 139*, 352– 402.

Xu, C., & LeFevre, J. A. (2016). Training young children on sequential relations among numbers and spatial decomposition: Differential transfer to number line and mental transformation tasks. *Developmental Psychology, 52*(6), 854.

CONNECTING MATHEMATICS LEARNING TO LEARNING ABOUT STRUCTURAL RACISM IN THE UNITED STATES

Eva Thanheiser and Molly Robinson

Portland State University

In this paper we examine the experience of mathematics majors enrolled in a course focused on examining mathematical concepts while learning about structural racism in the United States. Tasks were designed/modified to address both the deep conceptual understanding of mathematics in context (such as the meaning of ratio, rate, slope on a graph) as well as an understanding of systemic racism in the United States. Students developed a deeper understanding of mathematics in context, a deeper understanding of structural racism, and a broader understanding of what mathematics is.

INTRODUCTION & FOCUS OF THE PAPER

There are long-standing and ongoing calls for making mathematics meaningful, relevant, and applicable outside the classroom at all levels of education including the university classroom (Consortium for Mathematics and its Applications [COMAP] & Society for Industrial and Applied Mathematics [SIAM], 2016; National Council of Teachers of Mathematics [NCTM], 1989, 2000, 2014). To prepare democratic citizens requires going beyond "everyday" and "abstract" contexts and including the analysis of complex social and political issues in the mathematics classroom. In such classrooms mathematics becomes a tool for understanding, analysing, and changing the world (Freire, 1970; Freire & Macedo, 1987; Gutstein, 2006). This perspective, of mathematics as a tool for understanding and critiquing our socio-political world, along with other equity and justice concerns, has been endorsed by a number of professional organizations in the United States (AMTE, 2017; Larson, 2016; NCSM & TODOS, 2016; NCTM, 2000, 2014). In this paper we examine the experience of a group of mathematics majors who engaged in a course titled *The Mathematics of Racism* which addresses the calls to make math more meaningful and relevant. Students explored tasks that addressed both the deep conceptual understanding of mathematics in context (such as the meaning of ratio, rate, slope on a graph) as well as an understanding of systemic racism in the United States. The focus of this paper is on the following question: How did engagement with such tasks effect the students understanding of mathematics, understanding of structural racism, and experience in the course?

THEORETICAL FRAMING & BACKGROUND

Frames For Mathematics

Making mathematics meaningful, relevant, and applicable outside the classroom requires framing, or reframing, mathematics as a tool for making sense of the world and human activity. Lakoff explained "frames are mental structures that shape the way we see the world" (Lakoff, 2014, p. ix). Every word evokes at least one frame, but

2023. 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. 267-274). PME 46.

some words can evoke multiple frames. “Most frames are unconscious and have just developed naturally and haphazardly and have come into the public’s mind through common use.” (Lakoff, 2006, p. 2). Author (blinded for review) lays out three interrelated frames for the word mathematics that are in use in the mathematics education community: **Frame 1:** Mathematics as an abstract body of knowledge/ideas, the organization of that into systems and structures, and a set of methods for reaching conclusions. **Frame 2:** Mathematics as contextual, ever present, as a lens or language to make sense of the world. **Frame 3:** Mathematics as a verb (not a noun), a human activity, part of one’s identity. Each frame has a different implication for what is cantered in a mathematics classroom. Of note is that Frames 2 and 3 are both necessarily contextual; Frame 2 views math as sense-making in context, while Frame 3 views math as a contextual human activity.

Definition of Structural/Institutional Racism

Structural racism in the United States has been defined by Bailey et al. (2017) as “the totality of ways in which societies foster racial discrimination through mutually reinforcing systems of housing, education, employment, earnings, benefits, credit, media, health care, and criminal justice,” (p. 1453). It has also been defined by Lawerence and Keleher (2004) as “

> the normalization and legitimization of an array of dynamics – historical, cultural, institutional and interpersonal – that routinely advantage whites while producing cumulative and chronic adverse outcomes for people of color. (p. 1).

Taken together, these “patterns and practices in turn reinforce discriminatory beliefs, values and distribution of resources,” (Bailey et al., 2017, p. 1453). Structural racism is often invisible (Bailey, 2017) because it is everywhere and thus hides white supremacy. Another word for structural racism is institutional racism, which has been defined as

> patterns, procedures, practices, and policies that operate within social institutions so as to consistently penalize, disadvantage, and exploit individuals who are members of nonwhite racial/ethnic groups” (Better, 2008, p. 11)

Institutional racism is “unquestioned, self-perpetuating, and powerful.” (Carroll et al., 1975, p. 16). Structural/institutional racism (from now on referred to as structural racism in this paper) permeates all facets of life including higher education, and especially STEM (McGee, 2020), and is the root cause for the opportunity gap (Merolla & Jackson, 2019). Structural racism is often invisible to students, especially students with privilege. One goal of the *Mathematics of Racism* course is to make structural racism more visible.

Critical Race Theory & Underlying Assumptions

This paper draws on Critical Race Theory (CRT) (Ladson-Billings & Tate, 1995; Solórzano & Yosso, 2002; Tate, 1997) and its underlying assumptions that race is socially constructed (Solórzano & Yasso, 2002). Two definitions of racism guide our

work (aligned with Solórzano & Yasso): "the belief in the inherent superiority of one race over all others and thereby the right to dominance" (Lorde, 1992, p. 496) and "a system of ignorance, exploitation, and power used to oppress African-Americans, Latinos, Asians, Pacific Americans, American Indians and other people on the basis of ethnicity, culture, mannerisms, and color" (Marable, p. 5). Our methodology is informed by several of the CRT tenants developed in the field of education, including (a) centrality of racism; (b) commitment to social justice; and (c) interdisciplinary perspectives (Davis & Jett, 2019; Delgado & Stefancic, 2001; Solórzano & Bernal 2001), to examine systems and spaces for white supremacy using mathematics. We approach this paper with an understanding of the sociohistorical and political contexts of slavery in the US, white supremacy, and antiblackness (Bell, 2018; DiAngelo, 2018; D'ignazio & Klein, 2020; Kendall, 2021; Kendi, 2016; Love, 2019; Rothstein, 2017; Saad, 2020; Steele, 2011; Tatum, 2017; Wilkerson, 2020) which result in overrepresentation of BIPOC (Black, Indigenous, and People of Colour) in certain contexts, such as remedial math classes, and underrepresentation of BIPOC in others, such as advanced math classes. The underlying assumptions in this paper are that if things were fair, everyone would have the same opportunities, and the population across all contexts would represent the population at large.

Seeing/Concretizing Structural Racism Through Exploring Math.

When teaching with a focus on making sense of both structural racism and a mathematics concept the focus of the mathematics classroom expands from Author's (blinded for review) Frame 1 to Frames 2 and/or 3 in which the focus of the context is to better understand structural racism. The tasks implemented in this paper all have dual, interrelated goals: broadening the students' understanding of mathematics and developing students' understanding of structural racism in the United States.

METHODS USED

The study took place in a mathematics course titled The Mathematics of Racism for mathematics majors at a large urban state university. In this course, we used various contexts such as the Impact of the 1965 Voting Rights Act (proportional reasoning, ratio, rate, percent) on representation in Congress (based on Munter & Haines, 2022) and Majority & Power (based on Wolfe & Amidon, 2022).

Students were asked to examine the contexts with interrelated mathematics content and social and political context goals. In addition, we read the book Weapons of Math Destruction (O'Neil, 2016). The participants in the study were 12 students enrolled in the Mathematics of Racism, a university capstone course. As such all students were seniors (in their last term before graduating). Of the 12 students 10 were men and 2 were women. Of the 12 students 2 identified as Asian and 10 as white. Of the 10 one identified as Hispanic while the other 9 identified as non-Hispanic. The focus of this course was described as: We will explore how mathematics can be used to understand, explore, and investigate racial and social injustices in the United States. We live in a

society where mathematics is at the foundation of many injustices. We will use mathematics to explore and examine various topics that allow us to understand systemic racism in the United States. The first author was the main instructor of the course. The course was a 6-credit course that met for 4 hours each week for 10 weeks (40 contact hours) and had 2 extra hours for a community project. Students typically worked on each task for a two-week period. The first week they examined the task and the mathematics, the second they explored further and made predictions. Data sources for the study are video recorded interviews conducted in the first week of class with each student and a video recorded final presentation presented by each student during the last week of class. Questions/Prompts for the interview included: What is math? And How would you define structural racism? Can you give an example? Sample final presentation prompts included: I used to think math is … Now I think math is …. And I used to think structural racism is … Now I think structural racism is … In addition, all class activity and student artifacts (Google Slides) were recorded.

Data was analysed initially by two researchers independently identifying themes for student responses, meeting to discuss and agree on a joint set of themes, returning to individually code by agreed on themes, meeting to discuss coding. All disagreements were resolved through discussion.

RESULTS

The students view of mathematics

The students' views of mathematics changed from the beginning to the end of the course. In the beginning of the course 8 of the 12 students described mathematics as a rigid and well-defined set of courses (sequence of topics), objective/neutral, or for smart people. The other 4 described it as being about understanding number and quantity, deductive reasoning, or getting solutions. At the end of the course all 12 students stated that math can be applied to anything, is context based, can used for good, or misused. We share some sample responses in Tables 1 and 2.

Table 1: Student responses to *what is math*.

	Before the Course: What is math?
Andrew	Umm, I'd say it's like the study of logic and patterns.
Gunther	Math is the study of I mean I guess you could say it's the study of numbers, but I guess it's kind of focused on the study of quantities
Martha	Math is using numbers, symbols, and letters to explain and solve problems.

While the student's views of mathematics aligned with Frame 1 at the beginning of the term, it broadened towards Frames 2 and/or 3 at the end of the course. Broadening the view of mathematics majors is particularly important as they are going to be the ones who might create new mathematics that could potentially create harm or teach mathematics.

Table 2: Student responses to I used to think math is … now I think math is …

	After the course: I used to think math is … Now I think math is ….
Andrew	I used to think math was rigid and well defined. Initially my understanding of math was limited to calculus, geometry, algebra, etc. and there wasn't much more than that mathematically. That topics in math could only have specific uses. Now I think math is applicable to abstract things and has many interpretations. There are complexities that can be interpreted mathematically and often the mathematical approach provides an accurate description of an event or occurrence.
Gunther	I used to think math was carefully applied to real life in order to avoid getting obstructed by the roughness of the real world and causing harm from extrapolation. Now I think math is often carelessly harnessed and trusted implicitly.
Martha	I used to think math was the language of science and a neutral tool for solving problems. Now I think math is the language of everything and specific care must be taken to ensure that the models we make don't cause harm and unfairness.

The students view of structural racism

Table 3: Student responses to what is structural racism

	Before the Course: How would you define structural racism?
Luis	Structural racism exists throughout our government politics, our legal system.
Steven	I mean I the first thing comes up to me … involves different social classes. I don't know if that's true or not.
Thomas	It's like a legacy of racist policies, of like consciously racist policies that, yeah, yeah. And then, and then, it's you know it's the unconscious biases that people have on the cultural scale, like that everyone has because of that history of pervasive racism

Table 4: Student responses to I used to think structural racism is … now I think structural racism is

	After the course: I used to think structural racism is … Now I think structural racism is ….
Luis	I used to think structural/systemic racism was only seen within government institutions that continue to find ways to keep racism alive. Now I think structural/systemic racism is everywhere in my life. With the use of math models in private industry, law enforcement, and government these models are adding to the structural/systemic racism that already existed.
Steven	I used to think structural/systemic racism was discrimination and prejudice against different hieratical social class. Now I think structural/systemic racism is unconscious prejudice that's deeply rooted in our institutions, rules, practices or social customs, that are not beneficial to people of color.
Thomas	I used to think structural/systemic racism was real, but subtle and difficult to define or prove the existence of Now I think structural/systemic racism is a well-defined phenomenon for which there is ample data-based evidence.

…

With respect to structural racism the students' understanding grew from a rough understanding of the concept that was limited to some institutions to a broader understanding of it being everywhere and being able to give more and more specific examples (see Tables 3 and 4)

After the course, students had a more detailed understanding of structural racism which makes it more visible to them, both as mathematics majors and as participants in the United States.

What was this course about?

When asked at the end of the course what students thought the course was 11 out of 13 stated that it was about understanding systemic racism with math, the connection of math to our world, and how misunderstanding math can have huge impacts on our lives. The two who did not mention this focused on visualizations to communicate math, which was also a large part of the course.

CONCLUSIONS

The dual goals of this course was to broaden mathematics majors' view of what mathematics is and develop a lens for seeing structural racism in the United States. Both goals were achieved, thus showing that it is possible to teach math in a way that allows students to understand the world they live in, even at the college level, and broaden their understanding of what mathematics is.

REFERENCES

Association of Mathematics Teacher Educators. (2017). Standards for preparing teachers of mathematics. Retrieved from http://amte.net/standards

Bailey, Z. D., Krieger, N., Agénor, M., Graves, J., Linos, N., & Bassett, M. T. (2017). Structural racism and health inequities in the USA: evidence and interventions. The lancet, 389(10077), 1453-1463.

Bell, D. (2018). Faces at the bottom of the well: The permanence of racism. Hachette UK.

Berry, R., Conway, B., Lawler, B., and Staley, J. (2020). High school lessons to explore, understand, and respond to social injustice. Thousand Oaks, CA: Corwin Press/Reston, VA: National Council of Teachers of Mathematics.

Better, S. (2008). Institutional racism: A primer on theory and strategies for social change. Rowman & Littlefield (p. 11).

Carroll Massey, G., Vaughn Scott, M., & Dornbusch, S. M. (1975). Racism without racists: Institutional racism in urban schools. The Black Scholar, 7(3), 10-19. (page 16)

Consortium for Mathematics and its Applications (COMAP), & Society for Industrial and Applied Mathematics (SIAM). (2016). Guidelines for assessment & instruction in mathematical modeling education. Retrieved from http://www.siam.org/reports/gaimme.php

Conway, B., Id-Deen, L., Raygoza, M., Ruiz, A., Staley, J. & Thanheiser, E. (2022). Middle school lessons to explore, understand, and respond to social injustice. Thousand Oaks, CA: Corwin Press/Reston, VA: National Council of Teachers of Mathematics.

Davis, J., & Jett, C. C. (Eds.). (2019). Critical race theory in mathematics education. New York, NY: Routledge.

DiAngelo, R. (2018). White fragility: Why it's so hard for white people to talk about racism. Beacon Press.

D'ignazio, C., & Klein, L. F. (2020). Data feminism. MIT press.

Freire, P. (1970). Pedagogy of the oppressed. New York, NY: Seabury.

Freire, P., & Macedo, D. (1987). Literacy: Reading the word and the world. Westport, CT: Bergin & Garvey.

Gutstein, E. (2006). Reading and writing the world with mathematics: Toward a pedagogy for social justice. New York, NY: Routledge.

Kendall, M. (2021). Hood feminism: Notes from the women that a movement forgot. Penguin.

Kendi, I. X. (2016). Stamped from the beginning: The definitive history of racist ideas in America. Hachette UK.

Ladson-Billings, G., & Tate, W. F. (1995). Toward a critical race theory of education. Teachers College Record, 97(1), 47–68.

Larson, M. (2016). A renewed focus on access, equity, and empowerment. NCTM's President's Messages. Retrieved from http://www.nctm.org/News-and-Calendar/Messages-from-the-President/Archive/Matt-Larson/A-Renewed-Focus-on-Access,-Equity,-and-Empowerment/

Lawrence, K., & Keleher, T. (2004, November). Structural racism. In Race and public policy conference

Lorde, A. (1992). Age, race, class, and sex: Women redefining difference. In M. Andersen & P. Hill Collins (Eds.), Race, class, and gender: An anthology (pp. 495-502). Belmont, CA: Wadsworth.

Love, B. L. (2019). We want to do more than survive: Abolitionist teaching and the pursuit of educational freedom. Beacon Press.

Marable, M. (1992). Black America. Westfield, NJ: Open Media.

McGee, E. O. (2020). Interrogating structural racism in STEM higher education. Educational Researcher, 49(9), 633-644.

Merolla, D. M., & Jackson, O. (2019). Structural racism as the fundamental cause of the academic achievement gap. Sociology Compass, 13(6), e12696.

National Council of Supervisors of Mathematics, & TODOS: Mathematics for ALL. (2016). Mathematics education through the lens of social justice: Acknowledgment, actions, and accountability. Retrieved from http://www.todos-math.org/assets/docs2016/2016Enews/3.pospaper16_wtodos_8pp.pdf

National Council of Teachers of Mathematics. (1989). Curriculum and evaluation standards for school mathematics. Reston, VA: National Council of Teachers of Mathematics.

National Council of Teachers of Mathematics. (2000). Principles and standards for school mathematics. Reston, VA: National Council of Teachers of Mathematics.

National Council of Teachers of Mathematics. (2014). Principles to actions: Ensuring mathematical success for all. Reston, VA: National Council of Teachers of Mathematics.

Rothstein, R. (2017). The color of law: A forgotten history of how our government segregated America. Liveright Publishing.

Saad, L. F. (2020). Me and white supremacy: Combat racism, change the world, and become a good ancestor. Sourcebooks, Inc.

Solórzano, D. G., & Bernal, D. D. (2001). Examining transformational resistance through a critical race and latcrit theory framework: Chicana and Chicano students in an urban context. Urban Education, 36(3), 308–342.

Solórzano, D. G., & Yosso, T. J. (2002). Critical race methodology: Counter-storytelling as an analytical framework for education research. Qualitative inquiry, 8(1), 23-44.

Steele, C. M. (2011). Whistling Vivaldi: How stereotypes affect us and what we can do. WW Norton & Company.

Tate IV, W. F. (1997). Critical race theory and education: History, theory, and implications. Review of Research in Education, 22(1), 195–247.

Tatum, B. D. (2017). Why are all the Black kids sitting together in the cafeteria?: And other conversations about race. Hachette UK.

Wilkerson, I. (2020). Caste (Oprah's Book Club): The origins of our discontents. Random House

.

APPLETS AND PAPER & PENCIL TASKS AS RESOURCES FOR WORKING WITH MATHEMATICAL REPRESENTATIONS

Odelya Tzayada and Michal Tabach

Tel-Aviv University, Israel

Integrating applets into mathematics lessons is a challenge for elementary school teachers. The current study aims to evaluate to what extent applets chosen by the teachers functioned as resources, as compared to paper & pencil resources in mathematics lessons. In this study 42 teachers used, for the first time, applet activities in their lesson as well as paper & pencil tasks. The analysis focused on two mathematical aspects: the mathematical content and the required competence when engaging with representations. The findings show that teachers tended to prefer the applet activity over the paper & pencil task regarding representations. Furthermore, teachers preferred activities that move from the concrete to the abstract and not the opposite, regardless of the resource type.

INTRODUCTION AND THEORETICAL BACKGROUND

Visual representations can be classified into static or dynamic representations. In school, static representations include pictures in books, drawings, or sketches. Dynamic representations are presented in digital environments. Representations in digital environments were regarded as dynamic only when the user could interact with the objects on the screen (Moyer-Packenham & Bolyard, 2016). The user can manipulate the object on the screen in ways similar to manipulating concrete manipulatives. The user can slide, flip, and turn objects by using the computer interface almost as if they were concrete 3D objects. Moyer, Bolyard & Spikell (2002) emphasize that the added value of these dynamic representations lies in their allowing the user to derive meaning based on their actions. The virtual manipulative environments enable users to engage in an activity in a way that helps them discover and construct mathematical principles and relationships (ibid). Calder & Campbell (2016) add that the visual and dynamic elements of digital technologies change the way knowledge and understanding occur in learning.

Computer technologies have become powerful educational tools, leading many countries and individuals to turn them into solutions for educational needs (Meziane, et al.,1999). A wide variety of educational applets are available today over the internet (Kay, 2018), many of them being designed purposefully for education.

Mathematics resources can extend beyond the material object to incorporate human and cultural aspects, requiring teachers to shift from focusing on 'what' the resources are to 'how' they function in the classroom. (Adler, 2000). Previous studies have examined how applets functioned as resources in different mathematical contents, such as whole numbers (Loong, 2014), fractions (Reimer & Moyer, 2005), data and statistics (Suh, 2010), or geometry (Ng, Shi & Ting, 2020). Other studies asked to estimate

2023. 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. 275-282). PME 46.

applets' use relative to the mathematics competencies described in the curriculum, such as solving word problems (Lantz-Andersson et al., 2009) or higher-order thinking tasks (Lingefjard & Ghosh, 2022).

Two aspects concerning the teaching of mathematics in elementary school are discussed in the Israeli curriculum. The first describes the mathematical content that should be taught, while the second explains the required competencies and the corresponding learning opportunities that should be provided for students (INMPSC, 2006). The use of the applets and paper & pencil tasks as teaching resources is at the heart of this study. Hence, the study focuses on these two aspects.

RESEARCH QUESTIONS

In the context of elementary school teachers teaching a lesson for the first time in which they choose an applet and a paper & pencil task to implement in their class, we ask:

1) What connections may be found concerning the mathematical content between the applet activity and the paper & pencil task?

2) To what extent do the applet activity and the paper & pencil task provide opportunities to learn related to the representation competence?

METHODOLOGY

Forty-two elementary school teachers participated in a professional development program aimed at embedding mathematical applets in their teaching sequences. For all the teachers it was their first-time integrating applets into the teaching sequence.

Two complementary tools were used for data collection: (1) each teacher was asked to plan a lesson that used a mathematical applet along with the paper & pencil task, to enact the lesson in their class, and write a report about the implementation and a personal reflection, and (2) each teacher was asked to provide a short oral report on their experience in class, which was video recorded.

Data analysis

As mentioned above, the applet and the paper & pencil task were the mathematical resources that were used in the lessons. These resources and their mathematical role were examined relative to their contents and their required competence. Each lesson was classified into one of four *general* mathematical content areas: whole numbers, fractions, data and statistics, or geometry. To determine whether the mathematical goals of these resources implemented in the same lesson were similar, the *specific* mathematical content of the applet was compared to that of the paper & pencil task.

For example, the following applet and paper & pencil task were identified as similar, as the two resources both dealt with *whole numbers, including place value* (3rd grade). The activity in the applet was "How many balls (presented visually and dynamically on the screen) are on the surface?"; and the activity in the paper & pencil task was "Suggest different ways to decompose the number 853. (Use a variety of ways that are beyond the decimal decomposing)". The following resources were identified as

different. The activity in the applet was "Which part of the birds (presented on the screen standing on a wire) is green? Write as a fraction", which dealt with *Fraction as a part of quantity, operations in fractions* (4th grade); while the paper & pencil task was "Write a verbal story suitable for the exercise $\frac{2}{3} * 6$", which dealt with *Multiplying a whole number by a simple fraction* (6th grade, or 5th in case they interpreted multiplication as repeated addition).

Whenever representations supported the competence of transition from the concrete to the abstract or vice versa, the resource was classified according to the way it was utilized and the direction of the transition, as shown in Table 1.

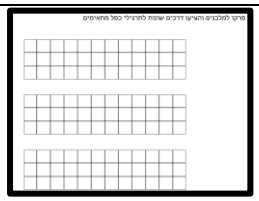
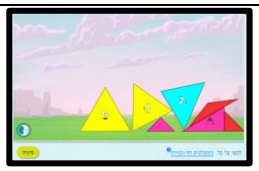
The required competence	Description	An example of applet activity	An example of a paper & pencil task
Transition from concrete to abstract	The resource - directs to convert visual (and dynamic) representations into abstractions	 *Tom covered half the cake with raspberries and ate 1/4 of it. What part of the cake did Tom eat?	 Suggest different ways to calculate the area based on dividing the rectangles into sub-areas.
Transition from abstract to concrete	The resource - directs to convert abstractions into visual (and dynamic) representations.	 *Press on all the acute triangles.	On five identical rectangles, I will ask the students to draw $\frac{1}{4}$ in different ways.

Table 1: The required competence regarding representations and its reflection in the resource.

Not all visual elements can be considered mathematical representations. When the visual element was aimed at motivating the user (Ben-Haim, Cohen & Tabach, 2019), it was not considered a mathematical representation. Also, in some cases a quantitative representation may express a change in the level of abstraction, but not a change between the concrete and the abstract. The analysis did not include such cases.

FINDINGS

Table 2 shows a distribution of the 42 lessons according to two aspects: the mathematical content and the required competence regarding representations. This forms the basis for answering the two research questions.

Table 2: The mathematical content and the required competence in the resources.

The mathematical content of the activities		The required competence regarding representations			
Specific	General	The applet activity		The paper & pencil task	
		From concrete to abstract	From abstract to concrete	From concrete to abstract	From abstract to concrete
Similar (32)	Whole numbers (14)*	7	1	3	2
	Fractions (9)	5	2	1	2
	Data and statistics (2)	2	--	1	--
	Geometry (7)	4	3	4	1
Different (10)	Whole numbers (3)	1	--	--	--
	Fractions (7)	5	2	1	2
	Data and statistics	--	--	--	--
	Geometry	--	--	--	--
Total (42)		24	8	10	7

*The numbers in brackets refer to the distribution of 42 lessons according to the mathematical content (specific /general). In addition, the numbers in the row(s) do not add up to 14. This is because the required competence could be reflected during the same lesson in more than one way and through more than one resource. Alternatively, in some lessons, it may not appear at all.

Research question 1: What connections may be found concerning the mathematical content between the applet activity and the paper & pencil task?

As Table 2 shows, in 32 lessons both resources referred to the same *specific* mathematical content. That is, for most teachers the choice of an applet from the database was accurate concerning the mathematical content of other resources in the lesson. This choice might have been made possible with the aid of the applets' classification system, which allowed teachers to identify suitable applets according to learning goals. Thus, the teachers could see beyond the interactive elements (visual and dynamic) of the applets despite the richness of the computerized environment.

When classifying all lessons into four general mathematical content areas, most of them dealt with whole numbers (17 lessons) and fractions (16 lessons). A small number of lessons dealt with geometry (7 lessons) and only two dealt with data and statistics. Contrary to expectations, geometry, whose visualization is an integral part of learning, was not prioritized when using applets. To make sense of these findings, in elementary mathematics curriculum there is a ratio of 1:6 between geometry lessons and all mathematics lessons. We found that a similar ratio of 7:42 was maintained in our study. We note that for the 42 participating teachers, it was the first time an applet was incorporated into the teaching sequence. Hence, it is possible that what led the teachers

to choose the applets was the representation's added value. When it comes to whole numbers and fractions, the applets serve as a significant visual support to the lack of resources in class in general (concrete manipulatives or textbooks). But in geometry, the textbooks include a wealth of relevant representations.

Research question 2: To what extent do the applet activity and the paper & pencil task provide opportunities to learn related to the representation competence?

This question was addressed by comparing between and within the applet and the paper & pencil task concerning identifying the transition between the concrete and the abstract and vice versa.

Comparing the two resources, the applet activity provided a more extensive response to the required competence relating representations compared to the paper & pencil task. In 31 lessons, the representations in the applet supported the required competence. For 30 of them, the transition was in one direction, either from concrete to abstract or from abstract to concrete. The paper & pencil task supported the required competence in only 17 lessons.

Concerning each part of the required competence separately (from concrete to abstract or from abstract to concrete): A total of 24 teachers selected an applet that provided an opportunity to learn the transition from concrete to abstract, while only ten paper & pencil tasks enabled this transition. It seems that the applet's environment provided a visual and dynamic richness that made it ideal for activities that provide a transition from the concrete to the abstract. However, in the transition from the abstract to the concrete, the applets (8) did not prove superior to the paper & pencil task (7). These findings raise questions, particularly since the computerized environment is characterized by accuracy in drawing, which might help in moving from the abstract to the concrete. Why did the teachers prefer not to use such a visually and dynamically rich environment over one that is implemented with paper & pencil? A possible explanation for this finding is that the teachers might not consider the accuracy of the drawing as significant. Any activity that enables the transition from the abstract to the concrete, even if not accurate, may have been seen as sufficient; a different explanation might be that teachers may have avoided giving their students activities that were visually and dynamically complex within the applet.

Comparing within each resource, the majority of teachers (24) preferred applets that supported students' transition from the concrete to the abstract, and a minority of them preferred applets that supported the transition from the abstract to the concrete (8). It is difficult to determine whether this ratio of 1:3 reflects the existing ratio of applets in the technological environment. For the paper & pencil tasks, teachers also preferred the transition from the concrete to the abstract, although the proportion was lower than in the applet. We can conclude that teachers preferred a transition from the concrete to the abstract, regardless of the resource that was in use.

Examining each mathematical content separately, we found that when using the applets for whole numbers and fractions, the teachers preferred the transition from the concrete to the abstract. For data and statistics, the applets served only the transition from concrete to abstract. This may be because this topic is deeply rooted in concrete, everyday life. In geometry, there was no difference between a transition from concrete to abstract (4) or abstract to concrete (3) when using the applets. It seems that for the teachers the applet environment did not offer any significant advantages over the visual environment in the textbooks. Teachers might not attribute much mediating value to the dynamic applets over paper & pencil tasks in geometry. In the paper & pencil tasks, the most noticeable difference was in fractions and geometry. In fractions, there were more tasks involving the transition from the abstract to concrete (4) and not from the concrete to abstract (2). In geometry, the activities from the concrete to abstract (4) were more than the abstract to concrete (1). This might be because the teachers might have tried to avoid giving their students tasks that required precise drawings.

SUMMARY AND DISCUSSION

Our study examined how applet activities and paper & pencil tasks were used as resources in mathematics lessons. This was done by focusing on two mathematical aspects: the mathematical content and the required competence when engaging with representations. The two research questions guide the discussion.

Research question 1: What connections may be found concerning the mathematical content between the applet activity and the paper & pencil task?

In this study, most teachers selected both resources so that the *specific* mathematical content in the applet activity was in coherence with the paper & pencil task. We assumed that this finding could be explained by the applet classification system the teachers used, which was accurate relative to the specific mathematical content. This finding is consistent with the study by Cabezuelo (2021). The researcher sheds light on the complexity involved in designing those systems. Indeed, it appears that the design solution of applet categorization, based on the domain's conceptual system, helped locate resources in the database.

As for the *general* mathematical content, only a sixth of the lessons dealt with geometry. This was contrary to our expectation that the visual richness of the technological platform would encourage teachers to conduct geometry lessons incorporating applets. Our findings are in line with Polly's study (2017). He found that only 9.28% of teachers reported enriching the curriculum with additional resources in geometry using internet resources. Hence, we may assume that curriculum contents available to teachers in the field of geometry (in textbooks, for example) are sufficient.

Research question 2: To what extent do the applet activity and the paper & pencil task provide opportunities to learn related to the representation competence?

For the transition from abstract to concrete, teachers showed no preference for a resource type. However, for the transition from concrete to abstract, teachers preferred applet activities over paper & pencil tasks. This finding may be explained based on Lee and Tan (2014). According to the researchers "virtual manipulative (is) narrowing (…) the cognitive gap between the concrete and pictorial representations" (p.108). They add that the ultimate goal, however, is to move from pictorial to abstract representations. Given that, our study may indicate teachers' pedagogical preference for dynamic over pictorial representations to scaffold students' ability in abstraction.

Within each resource, teachers preferred activities that deal with the transition from the concrete to the abstract, regardless of the resource used. The teachers' view, in our study, is in line with the Concrete-Pictorial-Abstract approach to mathematics teaching, which guides planning lessons that enable movement from the concrete to the abstract. According to this approach, to reach abstraction, one must first deal with the concrete (Lee & Tan, 2014).

Within each resource when using applets, teachers preferred the transition from the concrete to the abstract for three mathematics contents: whole numbers, fractions, and data & statistics. In paper & pencil tasks, teachers preferred the transition from concrete to abstract only in geometry. Also, they preferred the transition from the abstract to the concrete only in fractions. In their study, Divrik and Pilten (2021) examined students' performance in paper & pencil tasks on fractions. They found that in two fraction tasks, one dealing with the transition from abstract to concrete and the other dealing with the transition from concrete to abstract, the students achieved similar levels of success (90% and 88% respectively). These findings do not provide a sufficient explanation, and this question requires further investigation.

REFERENCES

Adler, J. (2000). Conceptualising resources as a theme for teacher education. *Journal of Mathematics Teacher Education, 3*(3), 205-224.

Ben-Haim, E., Cohen, A., & Tabach, M. (2019). Types of graphic interface design and their role in learning via mathematical applets at the elementary school. *Eleventh Congress of the European Society for Research in Mathematics Education*, Utrecht University, Utrecht, Netherlands.

Cabezuelo, S. A. (2021). Development of a classification system for Mathematical Logic. *Journal of Information Science, 47*(2), 143-160.

Calder, N., & Campbell, A. (2016). Using mathematical apps with reluctant learners. *Digital experiences in mathematics education, 2*(1), 50-69.

Divrik, R., & Pilten, P. (2021) Analysis of the Primary School 3rd Graders' Errors on Fractional Numbers in Terms of the Unit Fraction, Symbol, and Model .*Journal of Educational Theory and Practice Research, 7* (1), 62-73.

Israel National Mathematics Primary School Curriculum (INMPSC) (2006). Retrieved January 9 ,2023, from: https://meyda.education.gov.il/files/mazkirut_pedagogit/matematika/tochnyotlemud/mavo1.pdf

Kay, R. (2018). Creating a framework for selecting and evaluating educational apps. *Proceedings of the 12th International Technology, Education and Development Conference, Valencia, Spain*, 374–382.

Lantz-Andersson, A., Linderoth, J., & Saljo, R. (2009). What's the problem? Meaning making and learning to do mathematical word problems in the context of digital tools. *Instructional Science, 37*(4), 325-343.

Lee, N. H., & Tan, B. L. J. (2014). The role of virtual manipulatives on the concrete-pictorial-abstract approach in teaching primary mathematics. *Electronic Journal of Mathematics and Technology, 8*(2), 102-122.

Lingefjard, T., & Ghosh, J. (2022). Enhancing students mathematical thinking using applets. *Journal of Mathematics and Science Teacher, 2*(2).

Loong, E. Y. K. (2014). Fostering mathematical understanding through physical and virtual manipulatives. *Australian Mathematics Teacher, 70*(4), 3-10.

Meziane, F., Onn, C. W., Labadin, J., & Harris, R. (1999). Towards a framework for developing educational software. *Computers & Education, 30*(3/4), 237-274.

Moyer, P. S., Bolyard, J. J., & Spikell, M. A. (2002). What are virtual manipulatives? *Teaching children mathematics, 8*(6), 372-377.

Moyer-Packenham, P. S., & Bolyard, J. J. (2016). Revisiting the definition of a virtual manipulative. In Moyer-Packenham P. S. (Ed.) *International perspectives on teaching and learning mathematics with virtual manipulatives* (pp. 3-23). Springer.

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), 1-13.

Polly, D. (2017). Elementary school teachers' uses of mathematics curricular resources. *Journal of Curriculum Studies, 49*(2), 132-148.

Reimer, K., & Moyer, P. S. (2005). Third-graders learn about fractions using virtual manipulatives: A classroom study. *Journal of Computers in Mathematics and Science Teaching, 24*(1), 5-25.

Suh, J. (2010). Leveraging cognitive technology tools to expand opportunities for critical thinking in elementary mathematics. *Journal of Computers in Mathematics and Science Teaching, 29*(3), 289-302.

GROUNDING CHINESE NEW STANDARDS' FOCUS ON COUNTING-UNITS IN A CONSTRUCTIVIST, UNITS-AND-OPERATIONS MODEL

Ron Tzur[1], Rui Ding[2], Yunpeng Ma[2], Rongjin Huang[3], Bingqian Wei[1]

[1]University of Colorado Denver, [2]Northeast Normal University, [3]Middle Tennessee State University

In 2022, the Chinese Ministry of Education has released new Standards for primary and junior-high mathematics education. The Standards seem rooted in its authors' (expert) mathematical frame of reference, including an organization along a big idea of counting-units. Yet, this big idea opens the door for grounding the Standards in a constructivist, units-and-operations model that explains mathematical thinking and learning from the child's (inferred) frame of reference. In this theoretical paper, we offer brief overviews of the new Chinese Standards, including the big idea of counting-units, and of the constructivist (mental) units-and-operations model. Then, we propose how the latter can serve as a cognitive grounding of the former. Finally, we discuss potential benefits and importance of such grounding, as well as its likely challenges.

Mathematics education communities around the world (hence, PME) face a paramount issue: Can and how may research become relevant for practice at a national scale? We thus address the research problem: How may an empirically grounded, units-and-operations (constructivist) model guide choices made by curriculum designers, math teacher educators, and classroom teachers about (a) goals for students' learning and (b) instructional materials and methods to achieve them? Addressing such a problem in the Chinese context seems a conducive start, as it illustrates implementation of such choices at a very large national scale. In this paper we provide some preliminary insights into this great opportunity and formidable challenge.

As a theory of knowing and learning (not of teaching), constructivism can inform curriculum and teaching practices (Simon, 1995; Steffe, 1990). It helps ground mathematics teaching as it addresses two central questions that underlie research and practice: (a) what does it mean to know (or not) a mathematical concept and (b) how may a person come from not knowing to knowing it? We contend that such answers can underlie mindful, effective teaching aiming to accomplish mathematical standards, by explaining why selected instructional methods and materials may foster intended conceptual advances. Such answers are instrumental when, as it often happens, some or all students fail to make the intended advance despite their teachers' earnest efforts. Simply put, grounding mathematical standards in a constructivist theory can provide teachers with a lens through which to examine, and continually improve, their practice.

Our focus on the Chinese context is motivated by two lines of work. First is the plethora of recent mathematics education research by Chinese and international researchers.

2023. 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. 283-290). PME 46.

Chinese student outcomes on international comparative studies (OECD, 2020) seemed a major incentive for this growth in scholarly interest. Researchers thus undertook studies to examine aspects of Chinese students' learning of mathematics as well as of teaching practices used to achieve these outcomes (Ding et al., 2022).

We are also motivated by the recent commitment in China to reform and enhance its mathematics education, manifested through the publication and revision of its national *Standards* (Ministry of Education of the PRC, 2001, 2011). These Chinese Standards indicated a unique integration of reform foci (e.g., reasoning, critical thinking, problem solving) into a national system. The new, national *Standards* (Ministry of Education of the PRC, 2022) for mathematics in Chinese primary and junior-high schools further showed this continual enhancement effort. These new *Standards* included a major revision of primary school mathematics around big ideas, among them "counting-units" (Gong et al., 2022; Ma, 2022). Next, we turn to key aspects of this revision.

CHINESE NEW STANDARDS FOR PRIMARY SCHOOL

A major thrust for the recently revised *Standards* (2022) has been to create a structured integration of curricular contents by enhancing the mathematical integrity, consistency, and stages of learning themes in grades 1-9 (Ma, 2022). The *Standards'* authors endeavoured to find big ideas that can help integrate learning themes in four main fields: number and algebra, geometry, statistics and probability, and synthesis and practice (i.e., problem solving, project-based learning). They chose *counting-units* as one such organizing big idea. Using it enabled, for example, to revise six previous (scattered) themes in Number and Algebra into two themes organized through the lens of counting-units: (a) Numbers and Operations and (b) Quantitative Relationships.

The counting-units big idea seems rooted in the *Standards* authors' (expert) frame of reference. Ma (2022) noted that "operations in primary school are all number operations, including integer, decimal, and fraction operations. Numbers and operations are inseparable … Every subsequent [natural number] (+1) is added from 1 to get a new number, which contains the operation of addition" (pp. 35-36). When claiming that an expert's mathematical knowledge serves as the frame of reference, we underlie an issue: "For whom" does the big idea make sense? For example, Gong et al.'s (2022) characterisation of numerical operations does not seem (to us) to reflect an elementary student's frame of reference: "One is the borrowed definition. For example, we define 2 = 1 + 1 and 3 = 2 + 1 (*Peano's system of arithmetic axioms*), so that 1 + 1 = 2 and 2 + 1 = 3 …" (p. 47; italics added by this paper's authors). Such an expert's stance also seems a reason to claim that the meaning of all arithmetical operations "can be reduced to additive operations" (p. 45) and that "[t]he search for the arithmetic and algorithm of multiplication can, of course, be simply reduced to addition" (p. 48).

Most importantly, for us, is the underlying, expert's understanding that the operations are on numbers with the same counting-units: "The addition calculation of integers, decimals, and fractions can be understood as the addition of numbers in the same

counting-unit" (Ma, 2022, p. 36). Gong et al. (2022) further explicated the assertion that the essence of counting activities is to create counting-units:

> There are two main lines in the development of … counting activities: one is the generation and development of counting symbols, that is, from physical symbols to pictographic symbols, and then to abstract symbols; The second is the development of counting methods, that is, from "**one by one**" (non-carry system) to "group by group" (carry system, 进位制), from [a] counting symbol representing the same value at different places (non-place value system) to representing different values (place value system) … [It] can be seen that fractions are no less dependent on counting-units than integers. (pp. 46-47)

Another example of this assertion is given by Ma (2022) regarding the meaning and expression of numbers as abstract symbols of counting-units: "35 indicates three tens (TENS) and five ones (ONES), 3/5 denotes three 1/5s (fractional units), -35 denotes the amount opposite to 35, and so on." (p. 40).

To recap, the new Chinese *Standards* emphasise the big idea of counting-units as a useful, expert's lens for integrating numbers and operations into a whole that can serve as a basis for students to grasp substantial aspects of required mathematical knowledge. We emphasise the expert's frame of reference as an asset of what standards can and should do: provide aims for mathematics teaching. We now turn to a useful tool for explaining how 1-9 students may experience and come to grasp mathematics in their frame of reference – and thus for grounding Standards in a theory of learning.

A CONSTRUCTIVIST UNITS-AND-OPERATIONS MODEL

A constructivist research program, led by Steffe and colleagues (Norton, 2018; Steffe, 1992; Steffe & von Glasersfeld, 1985), has been creating models of children's mathematical thinking and learning. These researchers explicated that such models strive to explain mathematics from the child's frame of reference as inferred by the researchers. To distinguish between the two frames of reference, Steffe (2000) coined the terms first-order and second-order models. A *first-order model* consists of the mathematical ways one uses to explain their reality, whereas a *second-order model* consists of explanations of someone else's mathematical (experiential) reality.

A core premise of this research program is Piaget's (1985) construct of assimilation. It entails that any person can only interpret and act in their milieu by using mental structures – schemes – that are already available to them. Glasersfeld (1995) postulated that assimilatory schemes consist of three inseparable parts: a *situation* ("recognition template") that interprets perceptual and/or mental 'input' and sets a goal, a mental activity (i.e., *operation*) triggered to accomplish that goal, and a *result* brought about by that goal-directed activity. Using this three-part construct, models of children's mathematical thinking and learning focus on what is inferred to be the *child's* intentions (goal) and operations (activity) on units that the mental system constructs and on the effect(s) of those operations that the child notices and explains. The premise of assimilation underlies the "for whom" issue we emphasised above. For example, the phrase "35 indicates three tens (TENS) and five ones (ONES)" may be obvious to an

adult (or a child) who has constructed a scheme for coordinating these two types of units. However, for many children, such a phrase may only indicate an anticipation of a collection of thirty-five ONES, or a less advanced, retroactive anticipation after having counted 35 objects in a 1-to-1 correspondence, or a much less advanced anticipation of "35" being the result of rote counting from 1 to 35.

Units-and-operations models of children's mathematics have explained conceptual progressions from prenumerical, through numerical (additive and multiplicative), to fractional and algebraic reasoning (for a summary, see Tzur, 2019b). In this paper, we focus on three main characterisations that underlie those specific models. The first pertains to the nature of units (perceptual, figural, or abstract). The second pertains to the "size" of units (singletons or 1s), composite units (larger than and composed of 1s or other units), and compilations of composite units. The third pertains to the operations one uses with units (e.g., iterating, disembedding, distributing) and the levels of units one can coordinate upon assimilating a task (Norton et al., 2015).

For space reasons, we illustrate those three characterisations with an example of a child who was asked to build 5 towers of 3 (red) cubes each and 7 towers of 3 (blue) cubes each. Then, a researcher covered the first set (red) with a piece of paper on which "5 towers, 3 cubes each" is written and the second set with another, similarly written piece of paper and asked: How many red cubes are there? If only operating on *perceptual* units, the child would say, "I cannot tell unless lifting the paper." With permission to do so, the child will point to each cube while counting from 1 to 15. If operating on *figural* units, the child may leave the cubes covered and use right-hand fingers as substitute items (1s) for the hidden cubes (1-2-3; 4-5-6, etc.) while keeping track of the triplets (composite units) on their left-hand (1-2-3-4-5). A child who operates on *abstract* units may say: "1 tower is 3 cubes, … 5-is-15," likely coordinating accrual of both unit "sizes" – 1s and composite. The child may then be asked: How many more blue cubes are there than red cubes? The child may first find the total of 1s in each compilation and then subtract (e.g., 7x3=21; 5x3=15; 21-15=6), or first find the difference in composite units between compilations (subtracting 7-5=2) and then operate multiplicatively on the resulting 2 composite units of 3 cubes each (2x3=6). McClintock et al. (2011) and Wei (2022) postulated that a scheme coordinating those total-first and difference-first strategies, called *unit differentiation and selection* (UDS), provides a cognitive basis for the fundamental, distributive property.

Critically for the units-and-operations model, at issue is not mainly or just the fact the child obtained a correct answer to all tasks. Rather, it is the child's intentions and actions in terms of the nature, size, and operations used to solve it. Thus, for example, using this model leads to clearly distinguishing between additive and multiplicative reasoning. The former involves no change in units (e.g., 3 cubes + … + 3 cubes = 15 cubes). The latter involves operating on units by distributing items of one composite unit over items of another, very different composite unit to find a total of yet another type of unit (e.g., "3 cubes-per-tower" x "5 towers" = "15 singletons"). Explaining multiplication this way means coordinating at least two levels of units, whereas the

response to the difference between two compilations (7-5) likely involves three levels (1s, composite units, and compilations). Thus, whereas a two-level units coordination may be sufficient to operate with 10s and 1s (e.g., to consider 35 as 3 TENS and 5 ONES), operating simultaneously on three levels of units is needed to meaningfully operate with 100s, 1000s, and larger units. In fractions, similar characterisations (see Tzur, 2019a) include unit fractions as multiplicative relations – not just/mainly parts of whole (e.g., 1/5 is a unit that the whole is 5 times as much of it), composite fractions as dual multiplicative relations (e.g., 3/5 is a unit that is 3 times as much as 1/5), and fractions of fractions (e.g., 1/6 is 1/3 of 1/2 of the whole and 1/4 of that unit fraction is 1/24 of the whole – a 3-level unit coordination).

GROUNDING CHINESE NEW STANDARDS IN UNITS-AND-OPERATIONS MODEL

The units-and-operations model is rooted in the constructivist scheme theory (Glasersfeld, 1995). In it, the answer to what it means to know a concept is to *have an anticipation* of the relationship between a mental, goal-directed action – operation on/with some units – and the effects of that mental operation. For example, in the task about 7 and 5 towers (3 cubes each), constructing a UDS scheme entails having an anticipation that total-first and difference-first sequences of operations on composite units and 1s in them necessarily produce the same answer. Such an anticipation provides a cognitive basis for sensibly using the distributive property.

The units-and-operations model is situated in a constructivist stance on conceptual change. To explain how a person may advance to knowing a new (to them) concept, the theory builds on Piaget's (1985) premise of *reorganization*, which has profound consequences for teaching. According to this premise, a concept that someone does not yet know cannot be simply "transmitted" or "added-up" to what they already know. Rather, reorganization entails that all new knowledge is constructed through a person's active process, termed *reflective abstraction*, of mental changes to schemes that are already available to them. Simon et al. (2004) elaborated on it, postulating a mechanism termed *reflection on activity-effect relationships* (Ref*AER), so it could inform teaching. In a nutshell (for details, see Tzur, 2019a), the mental system uses two types of reflections and advances through two main stages, while beginning from an available scheme and reorganizing the anticipation of effects linked to previous goal-directed operations on units into a new anticipation. To illustrate how a units-and-operations model can provide grounding for the Chinese *Standards*, we use an example corresponding to those found in Ma (2022) and Gong et al.'s (2022) papers.

Tzur et al. (2013) postulated a 6-scheme conceptual progression in multiplicative reasoning; the UDS scheme discussed above is the third. Wei (2022) explained that the UDS scheme is a reorganization of the first two schemes, albeit in a reverse order: A person first finds the difference in composite units (second scheme, e.g., 7-5=2 towers) and then the total of 1s just in those two composite units (first scheme, e.g., 2*3=6 cubes). UDS can then be reorganized into a fourth scheme, *mixed-unit coordination*

(MUC), a basis for mindful operations in a place value, base ten system. Consider this task (bag = 10 oranges; box = 10 bags): School A has 2 boxes + 4 bags + 19 oranges; School B has 1 box + 16 bags + 11 oranges. Which school has fewer oranges and by how many? A first-order model of an adult likely includes coordinating multiplicative operations on composite units containing 10 bags of 10 oranges each. This seems to underlie an "obvious" solution that involves "counting units" of 1 in each of the given quantities separately (e.g., box=100, bag=10), and adding the 1s before comparing them (e.g., School A: 2*10*10+4*10+19=259; School B: 1*10*10 + 16*10 + 11=271).

However, using a new (validated) assessment of multiplicative reasoning, our data about Chinese primary students' solutions to such problems (forthcoming) showed that ~50% of them gave erroneous answers. In those students' frame of reference, it made sense to count (add) units of different magnitude (e.g., School A has 2+4+19=25, School B has 1+16+11=28). They likely assimilated the task into a scheme in which composite units (e.g., bags) are yet to be distinguished from compilations of those units (e.g., boxes), whereas operating multiplicatively on such different-magnitude units requires three-levels of coordination (e.g., a box is 10 bags of 10 oranges). That is, operating on and distinguishing units seems necessary for them to properly count units. Grounding *Standards* pertaining to relevant concepts (e.g., place value) would entail helping teachers develop a second order model that includes understanding why such erroneous answers make sense to the student and how instruction may foster reorganization of available schemes, so they construct a proper anticipation.

DISCUSSION: POTENTIAL BENEFITS, IMPORTANCE, AND LIKELY CHALLENGES

In this paper, we attempted to make a modest contribution to the potential impact that mathematics education research and theory may have on practice. Specifically, we examined how the Chinese new *Standards* (2022) may be linked with and grounded in a constructivist, units-and-operations model. Such grounding can strengthen linking of *Standards* content coherence with stages in students' mathematical development. We contend that such grounding can enhance large-scale efforts to reform practices at a national level. In this way, our paper also contributes to attempts of linking research in diverse social-cultural settings to benefit students' learning (see Ding et al., 2019).

We recognize three challenges that our approach introduces. First, Chinese colleagues in our team noted the use of "didactics" as a term used in their country in reference to cognitive development. However, as discussed in this paper, a key challenge they point to is how to promote the necessary distinction between students' and experts' frames of reference, among the authors of the *Standards* and ultimately among teachers. Teachers' construction of second-order models can be very difficult and require substantial, targeted professional development (PD). A second challenge is the additional work required of mathematics educators (e.g., in China) to both distinguish and explicitly link conceptual progression frameworks (see Steffe & Cobb, 1988; Steffe & Olive, 2010; Tzur, 2019b) with curricular sequencing proposed in the new

Standards (2022). Our examples of two schemes (UDS and MUC) in the progression of multiplicative reasoning give a hint at this direction. A third and perhaps the most serious of all challenges is the work required, with mathematics educators and teachers, to revise their practices to reflect what Tzur (2013) called *student-adaptive pedagogy*. Such a pedagogical approach embraces a mindset and practice of tailoring curricular goals and activities to students' conceptual whereabouts. To this end, substantial PD efforts will need to focus on providing teachers with instructional materials and methods that not only augment the *Standards* with conceptual underpinnings (e.g., of the counting-units big idea) but also let teachers go beyond "following" the examples of expert teachers to flexibly and mindfully foster learning as a reorganization of what students do know, that is, to have agency in implementing their practice.

REFERENCES

Ding, M., Wu, Y., Liu, Q., & Cai, J. (2022). Mathematics learning in Chinese contexts. *ZDM – Mathematics Education*, *54*(3), 477-496.

Ding, R., Jin, X. Z., Tzur, R., Wei, B., & Ma, Y. (2019). Constructivist teaching experiment: An evolutionary method for studying learning progressions. *Educational Science Research*, *7*, 54-60 [In Chinese].

Gong, Z., Shi, N., & Zhang, D. (2022). A new perspective on the revision of Compulsory Education Curriculum Standards in Mathematics: The consistency of the concept and operation of rational numbers. *Curriculum, Teaching Materials, and Methods*, *42*(6), 45-51 [in Chinese].

Ma, Y. (2022). Focusing on core concepts and implementing core competencies: An analysis of the content structuring of Compulsory Education Curriculum Stantards: Mathematics. *Curriculum, Teaching Materials, and Methods*, *42*(6), 35-44 [in Chinese].

McClintock, E., Tzur, R., Xin, Y. P., & Si, L. (2011). Engendering multiplicative reasoning in students with learning disabilities in mathematics: Sam's computer-assisted transition to anticipatory unit differentiation-and-selection. In L. R. Wiest & T. D. Lamberg (Eds.), *Proceedings of PME-NA 33rd annual meeting* (pp. 164-172). University of Nevada, Reno.

Ministry of Education of the People's Republic of China. (2001). *Mathematics Curriculum Standards for Compulsory Education*. Beijing Normal University.

Ministry of Education of the People's Republic of China. (2011). *Mathematics curriculum standards for compulsory education*. Beijing Normal University.

Ministry of Education of the People's Republic of China. (2022). *Compulsory education mathematics curriculum standards*. Beijing Normal University.

Norton, A. (2018). Frameworks for modeling students' mathematics. *The Journal of Mathematical Behavior*, *52*, 201-207.

Norton, A., Boyce, S., Ulrich, C., & Phillips, N. (2015). Students' units coordination activity: A cross-sectional analysis. *Journal of Mathematical Behavior*, *39*, 51-66.

OECD. (2020). *Benchmarking the performance of China's education system*. PISA, OECD Publishing. Retrieved from https://doi.org/10.1787/4ab33702-en.

Piaget, J. (1985). *The equilibration of cognitive structures: The central problem of intellectual development* (T. Brown & K. J. Thampy, Trans.). The University of Chicago.

Simon, M. A. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for Research in Mathematics Education*, *26*(2), 114-145.

Simon, M. A., Tzur, R., Heinz, K., & Kinzel, M. (2004). Explicating a mechanism for conceptual learning: Elaborating the construct of reflective abstraction. *Journal for Research in Mathematics Education*, *35*(3), 305-329.

Steffe, L. P. (1990). Mathematics curriculum design: A constructivist's perspective. In L. P. Steffe & T. Wood (Eds.), *Transforming children's mathematics education* (pp. 389-398). Lawrence Erlbaum.

Steffe, L. P. (1992). Schemes of action and operation involving composite units. *Learning and Individual Differences*, *4*(3), 259-309.

Steffe, L. P. (2000). Perspectives on issues concerning self, paideia, constraints, viability, and ethics. In L. P. Steffe & P. W. Thompson (Eds.), *Radical constructivism in action: Building on the pioneering work of Ernst von Glasersfeld* (pp. 91-102). Routledge.

Steffe, L. P., & Cobb, P. (1988). *Construction of arithmetical meanings and strategies*. Springer-Verlag.

Steffe, L. P., & Olive, J. (2010). *Children's fractional knowledge*. Springer.

Steffe, L. P., & von Glasersfeld, E. (1985). Helping children to conceive of number. *Recherches en Didactique des Mathematiques*, *6*(2/3), 269-303.

Tzur, R. (2013). Too often, these children are teaching-disabled, not learning-disabled. In *Proceedings of the 11th Annual Hawaii International Conference on Education*. Author (DVD).

Tzur, R. (2019a). Developing fractions as multiplicative relations: A model of cognitive reorganization. In A. Norton & M. W. Alibali (Eds.), *Constructing number: Merging perspectives from psychology and mathematics education* (pp. 163-191). Springer Nature.

Tzur, R. (2019b). Elementary conceptual progressions: Reality check + implications. In J. Novotná & H. Moraová (Eds.), *Proceedings of the international Symposium on Elementary Mathematics Teaching* (Vol. 1, pp. 29-40). Charles University.

Tzur, R., Johnson, H. L., McClintock, E., Kenney, R. H., Xin, Y. P., Si, L., Woodward, J., Hord, C., & Jin, X. (2013). Distinguishing schemes and tasks in children's development of multiplicative reasoning. *PNA*, *7*(3), 85-101.

von Glasersfeld, E. (1995). *Radical constructivism: A way of knowing and learning*. Falmer. (1995)

Wei, B. (2022). *Relating elementary teachers' MKT with their task selection for teaching multiplicative reasoning: A constructivist viewpoint* [Dissertation, University of Colorado Denver]. Denver, CO.

ANALYSING THE QUALITY OF ADVANCED MATHEMATICS LECTURES REGARDING THE PRESENTATION OF THEOREMS AND PROOFS – THE CASE OF REAL ANALYSIS LECTURES

Karyna Umgelter[1] and Sebastian Geisler[2]

[1]University of Hildesheim, [2]University of Potsdam, Germany

This paper reports a study investigating the quality of advanced mathematics lectures, in particular regarding the presentation of theorems and proofs. We compare the presentation of theorems and proofs given in two real analysis courses by two different lecturers using a structured observation tool. The results show, that informal parts of theorems and proofs were underrepresented in both lecture courses. Nevertheless, we could identify differences in the ways theorems and related proofs were presented with regard to different phases of the processes of proof construction. Thus, students might have developed different proving skills as well as different understanding of proof construction and methods of proving by both lecturers. Finally, we discuss implications for future investigations based on our results.

INTRODUCTION

Advanced mathematics lectures are challenging for many students. This could be one reasons why many first-year students drop out from their mathematics study programs or change to another subject (Geisler, 2020). Paoletti et al. (2018, p. 2) state that mathematics lectures "have been unsuccessful in promoting student learning". Nevertheless, lectures are still the most common teaching mode at many universities (e.g., Artemeva & Fox, 2011). Despite the critical responses to lectures, Fritze and Nordkvelle (2003, p. 328) say: "[T]he lecture survives, probably because it serves many functions not so well observed in the present research". Because lectures remain a dominant teaching format in advanced mathematics, more research analysing the quality of mathematics lectures is necessary. Actually, we have found only a few empirical studies concerning characteristics and quality of mathematics lectures (e.g., Viirman, 2014; 2021; Rach et al., 2016).

Weber (2004, p. 116) describes mathematics lectures as a "definition-theorem-proof" format. Moreover, mathematics lectures usually consist of "chalk talk" (Artemeva & Fox, 2011): a lecturer writes mathematical content on the board or a screen and makes some oral comments concerning this content. In this paper, we present results of an observation study concerning the presentation of theorems and proofs in mathematics lectures – as a possible quality criterion – supported by a structured observation protocol. In the following, we describe our theoretical framework for quality criteria of mathematics lectures as well as previous research regarding the presentation of theorems and proofs in mathematics lectures.

2023. 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. 291-298). PME 46.

THEORETICAL PERSPECTIVE

Quality criteria of mathematics lectures

There is no common shared definition of quality regarding advanced mathematics courses or university teaching at all. Kiendl-Wendner (2016) states that university quality can be characterized according to the following features: forms including macro level (whole university) and micro level (single courses), and dimensions including process quality (clearly defined processes with standardized procedures), result quality (achievement of the goals set) and structural quality (adequacy of resource allocation). Because German and international undergraduate mathematics courses usually consist of lectures (e.g., Pritchard, 2015; Artemeva & Fox, 2011), we take a closer look at the micro-level-quality of mathematics lectures and investigate their process quality.

We use a synthesis of two definitions to describe a lecture. Viirman (2021, p. 467) defines a lecture as "a teaching mode involving one teacher and a large group of students with communication mainly directed from the teacher to the students [...]". In addition, lecturers give their lectures usually "on a pre-announced topic […]" (Bergsten, 2007, p. 48).

Theoretical conceptualizations concerning the quality of mathematics lectures are scarce. Lamm et al. (2022) have already evaluated the quality of instruction in post-secondary mathematics courses. Because their framework is tailored towards community college mathematics courses, it seems less adequate for proof-based mathematics courses at university. Therefore, we consider frameworks with regard to quality of advanced mathematics lectures from Bergsten (2007) and Rach et al. (2016). For our investigation, we use a framework from Rach et al. (2016). According to Rach et al. (2016), the quality of mathematics lectures consists of two aspects: general criteria and mathematical criteria. General criteria include components such as learner orientation, cognitive activation, instructional efficiency as well as clarity and structure. As afore mentioned, most advanced mathematics lectures have a definition-theorem-proof format. Likewise, theorems and proofs play an important part in mathematics lectures (e.g., Paoletti et al., 2018) and are therefore also of major relevance for the quality of a mathematics lecture. Thus, mathematical quality criteria include the presentation of definitions, theorems and proofs. In this paper we decided to focus on presentation of theorems and proofs.

Presentation of theorems and proofs in mathematics lectures

Boero (1999) describes theorem production and proof construction as a process that consists of the following phases: production of a conjecture, formulation of the theorem statement, exploration of the content, selection and enchaining of arguments, organization of enchained arguments into a proof, and approaching a formal proof. Thus, processes of theorem production and proof construction consist of rather formal aspects (e.g., formulation of the theorem) as well as rather informal aspects (e.g., exploration of the content).

Davis & Hersh (1981) note, that in many mathematics lectures only rigorous theorems and related finished proofs are presented. Many researchers argue that presentation of only formal aspects of mathematics does not enhance students´ learning (e.g., Leron & Dubinsky, 1995). Indeed, according to previous studies, many students struggle to handle theorems and proofs (e.g., Stylianides et al., 2017). One of the reasons might be that lecturers invest not enough time in presentation of informal aspects of theorems and proofs like justification of selection and enchaining of arguments during proof construction or they don't highlight mathematical claims and methods that can be used in future proofs (Fukawa-Connelly, 2014). However, researchers could observe presentation of formal mathematics like theorems and proofs as well as presentation of informal mathematics like informal reasoning processes for proof construction in mathematics lectures (e.g., Fukawa-Connelly & Newton, 2014; Gabel & Dreyfus, 2017). Rach et al. (2016) found that lecturers present formal definitions almost correctly, whereby they pay less attention to formal argumentation. At the same time, the study from Rach et al. (2016) indicates underrepresentation of informal aspects of theorems and proofs like production of a conjecture (e.g., motivation of a theorem), exploration of arguments (e.g., collecting ideas to an informal proof) or significance and summary (e.g., highlighting of central decisions and giving an outlook). Using observations, Viirman (2014) indicates that some lecturers even do not use any theorems or proofs in first-semester mathematics courses. Moreover, students get too few opportunities to participate in advanced mathematics courses (Paoletti et al., 2018) and are not involved in construction of proofs during lectures (Alcock & Weber, 2005).

THE PRESENT STUDY

Research question

Empirical research regarding the characteristics of advanced mathematics lectures, especially "based on observations of actual lecturing" (Viirman, 2021, pp. 467), is rare. Our goal is to use a structured instrument to enable comparisons of mathematics lectures given by different lecturers on the same topic. In our ongoing project, we plan to observe a large number of mathematics lectures to be able to give an overview of quality of advanced mathematics lectures.

For this paper, we conducted a pilot study with two lecturers on the quality criteria "presentation of theorems and proofs" following Boero's (1999) phases of theorem production and proof construction. The main purpose of the study was to check whether it is possible to differentiate between lectures concerning this quality criteria using a structured observation tool. In particular, we want to answer the following research question:

Is it possible to identify differences with regard to the presentation of the phases of theorem and proof construction processes in advanced mathematics lectures given by two different lecturers and which differences can be found?

METHODOLOGY

The analysis of the study regulations in linear algebra and real analysis from several universities in Germany has shown that the topics in real analysis are quite similar across different universities. Therefore, we have decided to observe lectures in real analysis especially to the very common topics sequences and series. Typical for these topics are definitions for sequence, series and convergence as well as convergence rules like the Bolzano-Weierstrass theorem.

To ensure comparisons between different real analysis courses and to make our investigation systematic, we decided to use a standardized observation protocol as a measuring instrument. Rach et al. (2016) developed a structured observation protocol based on Boero´s (1999) framework. First studies from Rach et al. (2016) have shown that their observation tool is able to reliably collect data with regard to presentation of formal and informal parts of theorems and proofs in advanced mathematics lectures. We have merely adapted it slightly for our own research.

Explanation: The statement (e.g., theorem, lemma, proposition, corollary) is named, noted and explained. This category has to be chosen when the result of category *Production of a contention* is formulated, stated in writing and explained. Arguments or proofs should not be coded here.	
	presented well – The claim is made visible for students (e.g., written on the blackboard) and no terms are used that were not previously known. The claim is explained, e.g., by clarifying terms, and what has to be proven is highlighted.
	presented – The claim is made visible for students (e.g., written on the blackboard), but used terms or notations are not previously known for students.
	presented poorly – The formulated claim is not complete, mathematically wrong or not what is to be proved.
	not presented – The claim is not formulated.

Figure 1: Observation protocol for the category *Giving of a formal theorem* (following Rach et al., 2016).

The standardized observation protocol describes the presentation of theorems and proofs in five categories that correspond to the phases proposed by Boero (1999): *Production of a conjecture* (presentation of a problem and motivation of a theorem to solve it), *Giving of a formal theorem* (formulation of a formal theorem), *Exploration of arguments* (preliminary considerations to prepare a formal proof), *Organization of arguments* (formal proving of the theorem), *Significance and summary* (highlighting central decisions and perspectives, e.g., giving an outlook on other theorems that can be proved in a similar way). The categories *Giving of a formal theorem* and *Organization of arguments* cover formal aspects, the other three categories cover rather informal aspects of theorems and proofs. Each category can be evaluated in four grades: 1 – not presented, 2 – presented poorly, 3 – presented, and 4 –presented well. For every grade of evaluation in each category there is an exact coding description in the observation protocol. Figure 1 shows an example how the category *Giving of a formal theorem* must be coded. According to the observation protocol, every presented theorem and related proof in the lecture, whether presented only verbal or written on the board, has to be coded in all five categories. For analysing presentation of theorems and proofs, we viewed written and verbal statements as a unit.

In this paper, we present the results of the observation of two video recorded lecture courses presented by two different lecturers (we call them *Lecturer A* and *Lecturer B*) from the same large public university in Germany and the same mathematics programs. Both lecture courses were designed for pure mathematics students and upper secondary pre-service teachers. As previously mentioned, we coded only theorems and related proofs that are part of the topics sequences and series. Because both lecture courses were video recorded, we have had a good opportunity to stop or to repeat the recordings and to make some notes. Moreover, to be sure that the observation protocol is reliable, we asked another mathematics education researcher to code presented theorems and proofs in one lecture by each lecturer. We calculated the Spearman correlation coefficient in order to check the interrater reliability. The correlation of $\rho = .68$, indicates a medium but sufficient correlation between both codings.

First of all, we analysed which theorems and related proofs were presented by the lecturers. We noticed that there are some similar theorems and proofs presented by both lecturers and built two groups: common theorems and proofs (presented by both lecturers), and additional theorems and proofs (presented by only one lecturer). We expect, that those so-called common theorems and proofs are canonical for the topics sequences and series and those so-called additional theorems and proofs are less canonical. Moreover, both lecturers gave “chalk talk” lectures.

RESULTS

Lecturer A presented only common theorems and related proofs in his lecture course. *Lecturer B* presented common theorems and proofs as well as additional theorems and proofs.

Table 1: Comparison of the observed lectures (means and standard derivations of the coded categories, ratings from *1 – not presented* to *4 – presented well*).

	Lecturer A common (n = 13)		*Lecturer B common (n = 13)*		*Lecturer B additional (n = 22)*	
	M	*SD*	*M*	*SD*	*M*	*SD*
Production of a conjecture	2.55	1.07	1.31	0.60	1.14	0.45
Giving of a formal theorem	4.00	0.00	3.54	0.84	3.55	0.83
Exploration of arguments	3.37	0.48	2.24	0.69	1.60	0.93
Organization of arguments	4.00	0.00	3.31	1.26	2.37	1.49
Significance and summary	2.46	1.30	1.85	1.29	1.41	0.83

We analysed 13 common theorems and related proofs presented by both lecturers as well as 22 additional theorems and related proofs presented by *Lecturer B (see* Table 1). The values in columns *Lecturer A common* and *Lecturer B common* include only the results concerning common theorems and proofs presented by *Lecturer A* and *Lecturer B*. *Lecturer B additional* includes results concerning additional theorems and proofs presented by *Lecturer B*.

The results suggests that both lecturers achieve the best ratings in categories *Giving of a formal theorem* and *Organization of arguments*. All formal theorems and proofs in both lectures mostly were presented correctly. According to the ratings of both lecturers, the motivation for presented theorems (category *Production of a conjecture*) was weakly or not presented at all in both courses.

First, we compared the results of presentation of common theorems and proofs given by both lecturers using Wilcoxon tests, revealing that the differences concerning *Production of a conjecture* and *Exploration of arguments* are significant ($p < .05$) and those regarding *Giving of a formal theorem* are weakly significant ($p < .10$).

Second, we compared the ratings for presentation of common and additional theorems and proofs of *Lecturer B* using the Mann-Whitney-U-Test. According to the results, only the differences regarding *Exploration of arguments* are significant ($p < .05$) and those regarding *Organization of arguments* are weakly significant ($p < .10$). Overall, *Lecturer B* achieves for presentation of additional theorems and proofs lower ratings than for presentation of common theorems and proofs. According to the rating for the category *Organization of arguments*, there was some incompleteness regarding the organization of arguments to formal proofs of additional theorems. Moreover, *Lecturer B* achieves in the category *Exploration of arguments* low ratings regarding the presentation of additional theorems and proofs. Thus, both in additional and common theorems and proofs *Lecturer B* presents informal aspects only seldom.

CONCLUSION AND DISCUSSION

We could observe that both lecturers gave definition-theorem-proof (Weber, 2004) real analysis lectures in typical “chalk talk” style (Artemeva & Fox, 2011). Regarding presentation of common theorems and proofs, both lecturers pay much attention to presentation of formal aspects (e.g., *Giving of a formal theorem*) but it seems that they do not invest much time in presentation of informal aspects of theorems and proofs. These results are in line with Fukawa-Connelly (2014) and the observations from Rach et al. (2016).

Furthermore, we analysed presentation of additional theorems and proofs given by *Lecturer B*. It seems that *Lecturer B* pays less attention to presentation of formal and informal aspects of additional theorems and proofs then of common theorems and proofs. Furthermore, it is noticeable that the results in the categories *Exploration of arguments* as well as *Organisation of arguments* differ to a large extend concerning presentation of common and additional theorems and proofs. An explanation for these results is that *Lecturer B* presented some additional theorems completely without related proofs.

According to these results, we expect that although *Lecturer B* presented more theorems and proofs in his lectures, students might have viewed more ways and methods concerning production and motivation of theorems as well as construction of proofs in the course given by *Lecturer A* because *Lecturer A* presented more informal aspects of theorems and proofs in his lecture course than *Lecturer B*. Summarizing,

students taught by different lecturers probably develop different understanding and sense of proof construction and methods of proving even if they attend a course in the same university and study under the same study regulations in the same study program. A study that investigates the quality of presentation of theorems and proofs in lectures *and* students' actual developed skills concerning proof construction could confirm these assumptions.

There are some limitations in our study. The sample of our study consists of only two different lecture courses. To substantiate our results, a larger sample is necessary. In addition, the correlation between our coding and coding of the second coder is only medium. Therefore, a revision of the observation protocol prior to further investigations is necessary. Nevertheless, the use of a structured observation protocol enabled a structured observation of mathematics lectures and a first comparison of their quality regarding the presentation of theorems and proofs. We could identify differences in the presentation of formal and informal aspects of theorems and proofs by both lecturers as well as presentation of formal and informal parts of common and additional theorems and proofs by *Lecturer B*.

The next step of our ongoing research is to extend our observation protocol to general criteria. Future observations supported by an extended structured observation protocol should help to generate a broader picture regarding the quality of mathematics lectures at German universities. We expect that future studies could help us identifying possible connections between the quality of mathematics lectures and performance of students in their mathematics programs.

REFERENCES

Alcock, P., & Weber, K. (2005). Proof validation in real analysis: Inferring and checking warrants. *The Journal of Mathematical Behavior, 24*(2), 125-134.

Artemeva, N., & Fox, J. (2011). The writing´s on the board: The global and the local in teaching undergraduate mathematics through chalk talk. *Written Communication, 28*(4), 345-379.

Bergsten, C. (2007). Investigating Quality of Undergraduate Mathematics Lectures. *Mathematics Education Research Journal, 19*(3), 48-72.

Boero, P. (1999). Argumentation and mathematical proof: A complex, productive, unavoidable relationship in mathematics and mathematics education. *International Newsletter on the Teaching and Learning of Mathematical Proof, 7*(8).

Davis, P.J., & Hersh, R. (1981). *The mathematical experience*. Birkhäuser.

Fritze, Y., & Nordkvelle, Y.T. (2003). Comparing lectures: effects of the technological context of a studio. *Education and Information Technologies, 8*(4), 327-343.

Fukawa-Connelly, T.P. (2014). Using Toulmin analysis to analyse an instructor´s proof Attitudes toward Mathematics. *International Journal of Mathematics Education in Science and Technology, 45*(1), 75-88.

Fukawa-Connelly, T.P., & Newton, C. (2014). Analyzing the teaching of advanced mathematics courses via the enacted example space. *Educational Studies in Mathematics, 87*, 323-349.

Gabel, M., & Dreyfus, T. (2017). Affecting the flow of a proof by creating presence – A case study in number theory. *Educational Studies in Mathematics, 96*(2), 187-205.

Geisler, S. (2020). Early Dropout from University Mathematics: The Role of Student´s Attitudes towards Mathematics. In M. Inprasitha, N. Changsri, & N. Boonsena (Eds.), *Proc. 44th conf. of the int. group for the psychology of mathematics education: Interim Vol.* (pp. 189-198). PME.

Kiendl-Wendner, D. (2016). Die Qualität der Hochschullehre und deren Messung. In Steirische Hochschulkonferenz (Eds.), Qualität in Studium und Lehre. Kompetenz- und Wissensmanagement im steirischen Hochschulraum (pp. 243-262). Springer Fachmedien.

Lamm, R., Mesa, V., Kohli, N., Duranczyk, I., Watkins, L., & Ström, A. (2022). Construct Validation of an Instrument That Assesses Quality of Instruction in Community College Algebra. *Community College Journal of Research und Practice*. https://doi.org/10.1080/10668926.2022.2132430.

Leron, U., & Dubinsky, E. (1995). An abstract algebra story. *The American Mathematical Monthly*, 102(3), 227-242.

Paoletti, T, Krupnik, V., Papadopoulos, D., Olsen, J., Fukawa-Connelly, T., & Weber, K. (2018). Teacher questioning and invitations to participate in advanced mathematics lectures. *Educational Studies in Mathematics, 98*(1), 1-17.

Pritchard, D. (2015). Lectures and transitions: from bottles to bonfires? In A.C. Croft, M.J. Grove, J. Kyle, & D.A. Lawson (Eds.), *Transition in Undergraduate Mathematics Education* (pp. 57-69). University of Birmingham.

Rach, S., Siebert, U., & Heinze, A. (2016). Operationalisierung und empirische Erprobung von Qualitätskriterien für mathematische Lehrveranstaltungen in der Studieneingangsphase. In A. Hoppenbrock, R. Biehler, R. Hochmuth, & H.-G. Rück (Eds.), *Lehren und Lernen von Mathematik in der Studieneingangsphase. Herausforderungen und Lösungsansätze* (pp. 1-13). Springer Spektrum.

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 (Eds.), *Handbook of research in mathematics education* (pp. 237-266). NCTM.

Viirman, O. (2014). The functions of function discourse – university mathematics teaching from a commognitive standpoint. *International Journal of Mathematical Education in Science and Technology, 45*(4), 512-527.

Viirman, O. (2021). University Mathematics Lecturing as Modelling Mathematical Discourse. *International Journal of Research in Undergraduate Mathematics Education, 7*(4), 466-489.

Weber, K. (2004). Traditional instruction in advanced mathematics courses: a case study of one professor´s lectures and proofs in an introductory real analysis course. *Journal of Mathematical Behavior, 23*, 115-133.

STRATEGY USE IN NUMBER LINE ESTIMATIONS OF FRACTIONS – AN EXPLORATORY STUDY IN SEARCH FOR ADAPTIVE EXPERTISE

Wim Van Dooren

Centre for Instructional Psychology and Technology, KU Leuven, Belgium

The current study aims to provide a first exploration of the strategy that learners employ when they estimate fractions on a number line, and the extent to which they adapt their choice of strategies to features of the items, i.e. the specificities of the fractions that need to be estimated. In individual interviews 69 sixth graders estimated 18 fractions on a number line. Their accuracy and the strategies employed were registered. Results showed that learners with higher mathematical ability used a greater variety of strategies, and that – regardless of mathematical ability – a greater variety of strategies used across the item set results in greater accuracy. Higher ability learners also use the more accurate strategies more often, and there are indications that they adapt their strategy use more to specific item features.

THEORETICAL AND EMPIRICAL BACKGROUND

In the last two decades, research has amply shown that learners struggle to understand rational numbers, and particularly more than natural numbers. Even adults seem to struggle with a number of aspects of the rational number concept (Vamvakoussi & Vosniadou, 2004; Van Hoof et al., 2017). One of the main obstacles that research refers to is that rational numbers essentially differ from natural numbers, and that the prior knowledge of natural numbers – which is acquired first, and is extensively practiced – interferes with the learning of and reasoning about rational numbers. We exemplarily elaborate on the understanding of the magnitude of rational numbers, as this is also the focus of the current paper: Numerous studies (e.g., Vamvakoussi et al., 2018) have shown that the magnitude of natural numbers is understood better than that of rational numbers. For instance, when comparing 0.53 and 0.7, learners may judge that 0.53 is larger than 0.7 because it is longer (a technique that would work for natural numbers), or that 8/13 is larger than 4/5 because 8 and 13 are larger than 4 and 5.

A good understanding of the magnitude of numbers is considered important for the later development of mathematics. Generally speaking, scholars (e.g., Booth & Siegler, 2006) assume that learners mentally represent numbers as ordered on a mental number line from small to large, and a better mental representation of numerical magnitudes is predictive of later mathematical achievement. Research on numerical magnitude understanding often uses number line estimation (NLE) tasks (Schneider et al., 2018), and most often the number-to-position task whereby a segment of the number line with a beginning and endpoint are given and a specific target number needs to be positioned

2023. 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. 299-306). PME 46.

on that number line. The accuracy of the estimates is then an indication of the numerical magnitude representation.

While NLE tasks have been extensively used to measure the understanding of natural numbers (e.g., Booth & Siegler, 2006; Siegler & Opfer, 2003), it is more recently also used with rational numbers (fractions and decimals) and in comparing natural and rational numbers. For instance, Iulcano and Butterworth (2011) found that both children and adults made less accurate estimates of fractions than of natural numbers and decimals. The current paper therefore takes a closer look at the estimation of fractions on the number line.

In recent years, researchers have come to understand that estimates on a number line do not occur directly, but are highly strategy based. Ashcraft & Moore (2012), for instance, saw that accuracy is higher for numbers close to the endpoints and to the midpoint, suggesting that learners use benchmarks. Peeters et al. (2017) documented this strategy use and its role in estimation accuracy in detail.

Some studies also looked at the strategies that are employed when estimating fractions on the number line. Zhang et al. (2017) identified a segmenting strategy where the number line is divided into a number of equally large segments (often the denominator) and then counting the required number of segments (the numerator). They also identified a numerical transformation strategy, whereby a fraction is transformed in a decimal number, which is then mapped to a number line with benchmarks, of which the main ones are the endpoints, midpoint, and quarters. Zhang et al. (2017) also focused on inaccuracies and errors when implementing such strategies. Siegler & Thompson (2014) roughly distinguished the same strategies, but additionally focused on a strategy that focused on the endpoints (e.g. 95/98 being nearly 1).

RATIONALE AND GOAL OF THE CURRENT STUDY

While number line estimation with natural and rational numbers has shown the importance of a good magnitude understanding for later mathematical performance, and while the strategic aspects of number line estimation have been documented for both kinds of numbers as well, the importance of a good strategy choice in estimating fractions on a number line has not yet been investigated. Still, given the wide range of features of fractions, it may be that specific strategies are more suitable for estimating fractions, and other strategies for other fractions. This can be linked to the notion of adaptive strategy use, as elaborated by Verschaffel et al. (2009): Depending on the specific features of an item that needs to be solved, a certain strategy may be more suitable (more accurate and/or faster) than another strategy. Additionally, the most appropriate strategy may differ from learner to learner, as certain learners may be better at implementing a specific strategy than others.

The current study therefore has the goal to provide a first exploration of the way in which learners estimate fractions on a number line, and the extent to which learners adapt their choice of strategies to features of the items, i.e. the specificities of the fractions that need to be estimated. We decided to conduct this study with learners

from the 6th grade of primary school, as these learners already gained some understanding of fractions (that are taught gradually from the 2nd grade on), while still showing sufficient variation among learners. We also took into account that among these learners, there would be a range of mathematical abilities, and thus we could also explore whether the choice of strategies and the adaptation of these strategies to item features would differ across learners with different mathematical abilities.

METHOD

This study received approval of the SMEC-Ethical Committee of KU Leuven (G-2021-3036). A sample of 69 learners from 6th grade of primary school participated, after obtaining an informed consent from their parents. We received background information on their general mathematical competency, based on the standardized, curriculum-wide test system that is administered by schools to monitor students. It divides learners in four main levels (22 learners belonged to level A, 18 to level B, 15 to level C and 14 to level D) with A-level students being the highest achieving ones. This information was obtained from the participating schools.

All learners were individually interviewed. They solved a number-to-position number line estimation task, with number lines ranging from 0 to 1 and no additional benchmarks provided. The length of the number line was 200 mm. A set of 18 items (see below) was administered in a random order. Learners were asked to think aloud while conducting the estimation, and when necessary, the experimenter would ask further questions about the strategy used, such as “how did you do this estimation” or “can you tell which steps you took when making the estimation?” The interviews were audiotaped for later analysis.

Analysis

Accuracy. The accuracy of learners’ estimates was quantified by the most commonly used measure in the number line estimation literature, i.e. the percentage of absolute error (PAE). The fraction 2/11 is situated at 36.36 mm on the 200 mm number line. If a learner marks the fraction at 40mm, the PAE score is 1.82%, which is calculated as follows:

$$\text{PAE} = \frac{|40 - 36.36|}{200}\ x\ 100 = 1.82\%$$

Strategy. For every trial, the strategy that the learner has used was coded. In a first step, we eliminated all trials (8.6%) where a learner clearly did not attempt to estimate the actual fraction size (for instance, when a learner estimated 4/65 somewhere above 0.5, saying that (s)he estimated 65 on a 0 to 100 number line). In the remaining trials, we applied a 3 step coding scheme, thereby coding (1) whether learners initially transformed the given fraction (e.g. to another nearby fraction, to its exact decimal representation, to an approximate decimal representation), (2) the actual estimation strategy (benchmarking, segmenting, stepwise segmenting, and no strategy), and (3) whether a further correction happened in the finalisation stage.

Given length restrictions, this paper focuses on the second stage, i.e. the actual estimation strategy. The estimation strategies used in coding were defined as follows:

- **Segmenting**: Dividing the number line in a number of equal parts, which are dependent on the fraction that is being estimated (e.g., estimating 4/7 by dividing the number line in 7 equal parts and counting 4 of them, while dividing it in 5 equal parts to estimate 3/5)
- **Stepwise segmenting**: Dividing the number line in a number of equal parts, after which one or more of the parts are divided further in equal parts (e.g. estimating 2/9 by first dividing the line in 3 equal parts, and then further dividing the first part in 3 equal parts and counting 2 of them)
- **Benchmarking**: Halving the number line and when necessary halving it once more to obtain the 0.25, 0.5 and 0.75 values (or even halving further to obtain the octiles). These are then used as benchmarks to estimate the value of a given fraction. The decimal value of these benchmarks is present in students' reasoning, and unlike with (stepwise) segmenting, item characteristics to not meaningfully alter the procedure.
- **No strategy**: learners immediately mark the position of the target fraction, without previously using any additional marks (either physically or mentally)

The interrater reliability of this coding was checked for a subset of 180 trials, which provided a 88.6% correspondence between two independent raters.

Design of item set

In order to see adaptive strategy use, the choice of an adequate item set was of crucial importance. We designed an item set based on a rational task analysis and previous studies on strategy use in number line estimation involving fractions, thereby maximally trying to predict that certain items would be more easily estimated with a specific strategy rather than another one. Some fractions were close to the benchmarks and endpoints (0, 0.25, 0.5, 0.75, 1), assuming that a benchmark strategy would be more helpful them: *0:* 4/65, *0.25:* 3/11, 15/61, *0.5:* 29/63, *0.75:* 9/13, 45/58, *1:* 51/53. For some fractions with a small denominator being a prime number, we expected segmenting to be a helpful strategy: 2/11, 1/3, 3/5, 4/7. Other fractions had a small non-prime denominator, which would allow stepwise segmenting: 5/6, 3/14, 5/12, 8/15.

An important comment is that there is no one-to-one correspondence between fractions and these strategy categories. Fractions unavoidably have features that would make more than one strategy (but typically not all strategies) helpful. For instance, 4/7 may elicit segmenting the number line in 7 equal parts, but it may also be seen as a fraction close to 0.5.

MAIN RESULTS

Table 1 provides an overview of the frequency with which each strategy was used, as well as the median PAE for each strategy. Segmenting is by far the most popular

strategy across all items and participants, while stepwise segmenting occurs very rarely, and benchmarking and no strategy are in between. However, stepwise segmenting is substantially more accurate than the other strategies, and 'no strategy' is substantially less accurate. A one-way ANOVA showed a significant effect of strategy on accuracy, $F(3, 1130) = 16.01$, $p < .001$, but post hoc contrasts showed only a difference between 'no strategy' and the others.

A first, rather general, way to investigate whether adaptive strategy use is related to the accuracy of estimations is correlating the diversity in strategy use (i.e. the number of strategies used by one learner, varying between 1 and 4) and the median PAE of that

Strategy	Frequency	Median PAE
Segmenting	43%	3.17%
Stepwise segmenting	7%	1.93%
Benchmarking	28%	3.33%
No strategy	22%	4.54%
Total	100%	3.25%

Table 1: Frequency of strategy use and median PAE per strategy

learner. This correlation was *Spearman Rho* = -0.41, $p < .001$. This correlation indeed indicates that the more strategies a learner employs throughout data sets, the better the performance (i.e. the lower the PAE) is.

Second, learners with a higher mathematical ability generally use more strategies (*Spearman Rho* = 0.26, $p = .04$). This is also reflected in their performance: A-level learners had a median PAE of 2.73%, while this was 3.17%, 3.33% and 4.84% for B, C, and D level learners, respectively.

Combining both factors (learners' mathematical ability and the diversity in strategy use) in a single linear regression model ($F(2, 62) = 15.87$, $p < .001$) shows that 25.3% of the PAE can be explained by mathematical ability, while an additional 8.6% of the variance is added by including the diversity in strategy use. This implies that even within a level of mathematical ability, learners who use more strategies across all items generally are more accurate than learners who use fewer or even just one strategy.

Third, the explanation for the higher accuracy of higher mathematical ability students may not just be situated in the greater number of strategies that individual learners employ. It is possible that learners with a higher ability level make more use of the more accurate strategies. This can be checked with the data in Table 2.

This table first of all shows that the most accurate strategy (stepwise segmenting) is indeed used more often by A- and B-level learners, and the least accurate strategy (no strategy) is used more often by D-level learners. However, while overall segmenting

and benchmarking are almost equally accurate, it is remarkable that A-level learners use segmenting less often and benchmarking more often than the lower level groups. So, it may be that they employ it more often with items where this is indeed more appropriate, where the other groups still use segmenting for these items.

Table 2: Frequency of strategy use per mathematical ability level

	A	B	C	D	Total
Segmenting	38%	52%	53%	30%	43%
Stepwise segmenting	8%	11%	5%	2%	7%
Benchmarking	35%	25%	24%	23%	43%
No Strategy	19%	12%	19%	40%	22%

A fourth issue to explore is the following: Regardless of the diversity of strategies per student and how often they are employed, students with a higher mathematical ability may simply apply each of the strategies with greater precision. Table 3 gives an overview of the median PAEs for each strategy per ability level.

Table 3: Median PAE per strategy per mathematical ability level

	A	B	C	D
Segmenting	2.57%	3.30%	3.23%	3.22%
Stepwise segmenting	1.87%	2.50%	1.26%	9.64%
Benchmarking	2.83%	2.62%	3.47%	4.53%
No strategy	3.41%	4.56%	5.20%	5.70%

Once more this indeed seems the case, at least for some strategies, and trends are less pronounced than for the other analyses reported so far: A-level learners are more accurate when segmenting than the other learners, and the same goes for using no strategy. Further, A- and B-level learners are somewhat more accurate when benchmarking. (Note that he data for stepwise segmenting are not very reliable, as D-level learners almost never used this strategy.)

As a final step, we exemplarily focus on the strategy use and accuracy of one item. (A discussion on all items is impossible due to length restrictions). The fraction 5/12 was chosen as part of the item set as it would allow for stepwise segmenting (e.g., first halving, then going to fourths or sixths, and then possibly to twelfths). The mean PAE for this item is 3.33%. Stepwise segmenting indeed led to a substantially higher accuracy, with a PAE of 1.67%. Benchmarking (basically: using half as a benchmark, and then correcting because 5/12 is a bit lower than 0.5) was less accurate (PAE = 4.33%), and segmenting (directly splitting the number line in 12 equal segments) was even less accurate (PAE = 6.00%). There were only 9 learners (6 from the A-level, 3

from the B-level) who applied this stepwise segmenting, while benchmarking was most popular overall, and the least accurate strategy (segmenting) was most often used by C-level learners.

CONCLUSIONS AND DISCUSSION

The central goal of this study was to explore the strategies that 6th grade students would use when estimating fractions on the number line, whether this strategy choice would in some way relate to features of the items, and whether using different strategies (in relation to these item features) would result in a higher accuracy. We additionally explored whether students of different overall mathematical ability would differ in their strategy use.

First of all, we found differences in the accuracy of various strategies. Directly estimating without using any benchmarks or segments beforehand seems to be the least accurate, and in cases where learners used stepwise segmentation, this led to most accurate estimations. Another important finding was that there was a rather strong correlation between the number of strategies that an individual learner employed and his/her estimation accuracy, suggesting that it is indeed good to use more than just one strategy for the entire item set. Not surprisingly, learners of higher ability levels performed more accurately, but more importantly, the regression analysis showed that the use of a larger number of strategies across the item set was positively associated with accuracy, regardless of mathematical ability.

A further analysis showed that learners with higher ability levels indeed make more often use of the strategies that generally deliver more accurate estimates (particularly the strategy of stepwise segmenting), while it was particularly notable that learners of the lowest ability level very often made a direct estimation without any visible or reported strategy, i.e. by directly pointing at a position on the number line. Higher ability learners are also somewhat more accurate in the execution of the various strategies, but these effects were less pronounced, so the better performance seems mainly due to the more frequent use of specific more accurate strategies and the use of a greater variety of strategies across the item set.

The item set was created so that specific features would make the use of specific strategies more easy or difficult, and consequently would lead to more or to less accurate estimations. We found exemplary evidence that learners – especially of higher mathematical ability levels – indeed too into consideration such item features in choosing among the range of available strategies. Such signs of adaptive expertise may be relevant for education, as also lower performing students might benefit from using this broader range of strategies, taking into account item features. Instruction that is focused on the various strategies to estimate fractions on the number line may also lead to an increased conceptual understanding of fractions as such, and therefore be valuable beyond the specific context of estimating the size of fractions on a number line.

ACKNOWLEDGMENT

The authors want to thank Mathise Hofmans for her collaboration in the design, execution and data analysis for the current study.

REFERENCES

Ashcraft, M. H., & Moore, A. M. (2012). Cognitive processes of numerical estimation in children. *Journal of Experimental Child Psychology, 111*(2), 246-267.

Booth, J. L., & Siegler, R.S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology, 41*(1), 189-201.

Iuculano, T., & Butterworth, B. (2011). Understanding the real value of fractions and decimals. *The Quarterly Journal of Experimental Psychology, 64*(11), 2088-2098.

Peeters, D., Verschaffel, L., & Luwel, K. (2017). Benchmark-based strategies in whole number line estimation. *British Journal of Psychology, 108*(4), 668-686.

Schneider, M., Merz, S., Stricker, J., De Smedt, B., Torbeyns, J., Verschaffel, L., & Luwel, K. (2018). Associations of number line estimation with mathematical competence: A meta-analysis. *Society for Research in Child Development, 89*(5), 1467-1484.

Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity? *Psychological Science, 14*(3), 237-243.

Siegler, R. S., & Thompson A. C. (2014). Numerical landmarks are useful - except when they're not. *Journal of Experimental Child Psychology, 120,* 39-58).

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., Christou, K. P., & Vosniadou, S. (2018). Bridging psychological and educational research on rational number knowledge. *Journal of Numerical Cognition, 4*(1), 84-106.

Van Hoof, J., Vamvakoussi, X., Van Dooren, W., & Verschaffel, L. (2017). The transition from natural to rational number knowledge. *In acquisition of complex arithmetic skills and higher-order mathematics concepts* (pp. 101-123). London, England: Elsevier.

Verschaffel, L., Luwel, K., Torbeyns, J., & Van Dooren, W. (2009). Conceptualizing, investigating, and enhancing adaptive expertise in elementary mathematics education. *European Journal of Psychology of Education, 24*(3), 335-359.

Zhang, D., Stecker, P., & Beqiri, K. (2017). Strategies student with and without mathematics disabilities use when estimating fractions on number lines. *Learning Disability Quarterly, 40*(4), 225-236.

MAKING SENSE OF ZERO TO MAKE SENSE OF NEGATIVE NUMBERS

Joëlle Vlassis[1] and Isabelle Demonty[2]

[1]University of Luxembourg, [2]University of Liège

Few studies so far have examined the meaning attributed to zero by students and how this can affect the handling of negative numbers. The objective of this paper is to analyse the relationship between the understanding of zero and the ability to perform integer additions. A paper-and-pencil test was submitted to 166 grade 6 students who had not been taught about negative numbers, in order to analyse the meaning they attributed to zero in relation to their ability to perform integer additions correctly. The results show that these students had two main conceptions of zero: zero as "nothing" and zero as a "point on a number line". We also found that students with this latter conception of zero were significantly more likely to succeed in integer additions.

THEORETICAL BACKGROUND

Few studies have considered the question of zero in learning, and in particular the relationship between understanding zero and the ability to perform operations with negative numbers. However, to perform operations with these numbers it is important to be able to envisage the existence of numbers below zero. All too often, zero is regarded as "nothing" (Wheeler & Feghali, 1983), the "nil" below which no number is conceivable; this reflects the idea of "absolute zero" (Glaeser, 1981).

Like many mathematical symbols, zero has various different meanings. According to Volken (2000) and Toma (2008), there are two main uses of zero, which are both essential, but also slightly different. The first is as a way of marking an empty place in our positional notation system for numbers. Volken (2000) refers to this as a "meta sign", in that it indicates the absence of other signs. The zero in this context therefore seems to be of a different nature from the digits 1 to 9. The second use is as a number in its own right, in the form that we currently attribute to it, 0. Historically, this use took a long time to become established: zero was not a natural candidate to be a number, as numbers initially designated sets of objects (Toma, 2008). If there were no items to count, there was therefore no need to mention this "nothing", and even less need to give it a symbol (Ruttenberg-Rozen, 2018). The polysemy of zero makes it complex to understand and use for today's students at all stages of schooling (Levenson et al., 2007). It can also create obstacles in the extension of the natural number domain to the integers (Glaeser, 1981).

The purpose of this paper is to examine the different meanings attributed to zero by grade 6 students who have not yet learned about negative numbers, and to relate these meanings to their ability to solve integer addition problems.

2023. 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. 307-314). PME 46.

Zero in history

Historically, according to Toma (2008), the first trace of a zero was not found until around 400 BC, among the Babylonians, who used a special symbol to indicate empty space in their positional notation system. This sign can be considered the precursor of zero notation. However, Ruttenberg-Rozen (2018) notes that zero as we know it today in our base-10 positional system was developed in India from the 5th century AD onwards, when mathematicians began to use a small circle to represent the empty space in a number. The invention of this symbol by the Hindus made conceptual progress possible with regard to zero, whose status changed from that of a mere placeholder to that of a number less than one – a number in its own right (Ruttenberg-Rozen, 2018). Thus, in a treatise on astronomy in 628, Brahmagupta was able to define zero as the subtraction of a number from itself ($a - a = 0$). However, it was not until the 1600s that zero finally held an uncontested place and mathematics could further progress with the inclusion of this important number (Toma, 2008).

From the new understanding of zero as a number, other mathematical understandings were able to develop. Volken (2000) emphasises that the "zero" symbol was also the starting point for remarkable developments in arithmetic and algebra, resulting in the appearance of a new conception of numbers that was more abstract and more unified. Numbers gradually acquired autonomy from the pre-existing objects they had hitherto been deemed to represent. This abstraction made it possible to consider the numbers below zero, the negatives, which are not properties of sets of objects (Toma, 2008). Zero thus became "origin zero" (Glaeser, 1981). However, the introduction of negative quantities in the West was slow and difficult, due in particular to this ambiguity of zero, and many mathematicians throughout history have found it difficult to distinguish "origin zero" from "absolute zero" (Glaeser, 1981).

Zero in school education

What is zero? When asked, many students will reply that zero is "nothing" (Russell & Chernoff, 2011). Primary school teachers in pre-service education also appear frequently to use the words "zero" and "nothing" interchangeably or synonymously (Levenson et al., 2007).

Among other misconceptions about zero, it has also been observed that students believe that zero is not a number, or that it is only part of the symbol for the number ten, or that it "doesn't do anything" and can therefore be ignored (Russell & Chernoff, 2011). Some of this confusion can be attributed to the fact that many students think of zero as being "nothing". According to Levenson et al. (2007), this conception hinders effective teaching of the deep and complex structure of zero.

From the point of view of the relationship between zero and integer operations, Peled et al. (1989) carried out a study on addition and subtraction of integers among primary school students before they had learned about negative numbers. They examined the intuitive models used by students to perform these operations. The authors showed that the main mental model used by students was that of the number line, which is in fact

relatively abstract. However, some students, while referring to the number line model, did not necessarily make effective use of it. They displayed the misconception known as the "divided number line" (DNL): the numbers were seen as two symmetrical sets on either side of zero, which was often thought of as a barrier rather than a number. By contrast, other students were already showing the ability to take a unified view of the positive and negative numbers and zero as integers, in other words the conception of the "continuous number line" (CNL). These students performed calculations smoothly by going "to the right" for addition and "to the left" for subtraction, without having to create special partitioning rules to move past zero.

METHOD

The objective of this paper is to answer the following two research questions.

What meaning do grade 6 students, who have not been taught about negative numbers, attribute to zero? (RQ1)

To what extent does the meaning attributed to zero go hand in hand with the ability to perform integer additions correctly? (RQ2)

The analyses presented in this paper come from a larger study presented previously (Vlassis & Demonty, 2022). A total of 166 grade 6 students in 13 classes at eight primary schools in the Grand Duchy of Luxembourg took part in the study by completing a paper-and-pencil test. The students in our sample had not been taught about negative numbers, as this topic, as well as the meaning of zero, is not included in the Luxembourg primary curriculum (MENFP, 2011). The paper-and-pencil test was designed to be taken individually, and took approximately one hour to complete. The test had two parts: the first part consisted of decontextualised questions relating to integer additions and subtractions, the role of zero, the order of negative numbers, opposite numbers, and mental computations in addition and subtraction operations with more than two terms, and in subtractions with two terms where the compensation strategy would clearly be useful. The second part consisted of contextual problems, the solutions to which required operations similar to those in the first part. For the purposes of this study, only decontextualised questions relating to (1) integer additions and subtractions, and (2) the role of zero were analysed.

The two questions used for the analyses were as follows:

Question 1 (Q1.a and Q1.b)

a) What answer would you give to the problem 0 – 14 = ?
b) In 0 – 14 = what does the zero represent?

re than one answer may be correct. Tick the one answer that seems most appropriate to you.

1. The zero has no real value and could be removed
2. It represents a position on the number line
3. It represents a nil quantity

4. The zero is used to separate the positive and negative numbers
5. It represents emptiness

The five options presented in Question 1.b derive from students' misconceptions as described in the theoretical background:

- Option 1 refers to the "doesn't do anything" conception (Levenson et al. 2007);
- Options 2 and 4 both concern the "origin zero" (Glaeser, 1981), and are based either on the CNL (2) or on the DNL (4) (Peled et al., 1989);
- Options 3 and 5 correspond to "zero = nothing" which refers to the "absolute zero" conception (Glaeser, 1981).

Question 2 (Q2)

4 – 6 = □	–3 – 5 = □	–5 + 8 = □
8 + □ = 5	□ – 3 = –8	-12 – 8 = □
□ + 9 = 6	□ + 4= –6	4 +(–9) = □
5 – 10 = □	–7 + 4= □	7 + □= –3

In Question 2, some additional items with natural numbers were added to prevent students from deducing that all the answers to this question were negative numbers. These items show good internal consistency (Cronbach's alpha: 0.90) (Vlassis & Demonty, 2022).

RESULTS

The role of zero

In order to identify students' conception of zero (RQ1), we have analysed in Table 1 below the options chosen by students in Question 1.b.

Table 1: Percentage of choices made by students regarding the conception of zero (Q1.b)

Conception of zero	Choice	% of students
Nothing or emptiness ("nothing")	1	8
	3	11
	5	9
	more than one option	12
	Total	40
Number line (NL)	2	8
	4	28
	more than one option	7
	Total	43
Mixed conception		9
No answer given		8
Total (N=166)		100

In Table 1, the five options have been divided into two main categories: the "nothing" category consisting of Options 1, 3 and 5 and related to the "absolute zero" conception, and the "number line" category (NL), referring to the "origin zero" conception (Options 2 and 4). First, as can be seen from the results presented in Table 1, despite the instruction asking students to select a single option among the five, some of them ticked more than one. It should be noted that most of the students who ticked several options did so within the same category, thereby demonstrating some degree of consistency. Only 9% of students selected options in both categories ("Mixed conception"). Next, it is notable that the students' choices were evenly distributed between the two categories (40% for "nothing" versus 43% for "NL"). It is striking that even before any learning about negative numbers, 43% of the students already chose the NL conception (exclusively), implying an acceptance of numbers under zero. We use the term "implying" deliberately, because the students were not interviewed and we do not know what they actually thought on this subject. It is possible that the general context of the test, in which most of the questions concerned negative numbers, led to this choice being favoured by some students – a choice that they perhaps might not have made in a neutral context. Within this category, most of the students selected Option 4, referring to the DNL, indicating that their understanding of the integers and zero was not yet unified.

Finally, in the "nothing" category, the majority of choices refer to the "zero = nothing" conception, with 19% of the students choosing either Option 3 (zero = nil quantity) (11%) or Option 5 (zero = emptiness) (8%). A mere 8% of students ticked Option 1 only. It might be concluded that the idea of a zero that "doesn't do anything" was not particularly widespread among the students in the sample. However, closer analysis of the data (not presented in Table 1) shows that this conception was indeed present among the students. Among the 12% of students who selected "more than one option", 10% chose Option 1 in combination with Option 3 and/or 5. These students apparently thought that the idea of emptiness or nil quantity could be combined with the idea of "doesn't do anything".

Relations between conception of zero and operations with negatives

In order to answer RQ 2, we examined the results of the students relating to operations with negatives according to the conception of zero that they revealed. We thus related the results for Question 1.b with those for Questions 1.a and 2, hypothesising that students presenting an NL conception would succeed better in the operations with negatives than those presenting a conception labelled as "nothing+", this latter group including this time not only the main conception of "nothing" (40%), but also the "mixed conceptions" (9%) and students who failed to answer the question (8%). Table 2 below presents the answers given to Q1.a (0 – 14 =) according to the conception of zero (Q1.b).

Table 2: The relationship between conception of zero (Q1.b) and correct solution to the question 0 – 14 = (Q1.a)

Conception of zero (Q1.b)	Correct answers (%) (Q1.a)	Incorrect answers (%) (Q1.a)				No answer given (%) (Q1.a)	Total (%)
	–14	*14*	*-0.86/ –16/–6*	*0.86/86/ 6*	*0*		
NL conception	73	1	10	-	-	16	100 (N= 70)
"Nothing+" conception	49	13	4	2	2	28	100 (N= 96)

The results in Table 2 show that students with an NL conception of zero were significantly more likely (73%) to give the correct solution, –14, than students with a "nothing+" conception. The correlation of 0.24 between the results obtained in Questions 1.a and 2 confirms this link between the two variables. It is also worth noting that all these students, even when they were wrong, gave a negative answer (with the exception of one student who answered 14), and were much less likely to fail to answer (16%, against 28% of "nothing+" students). Forty-nine percent of the students with a "nothing+" conception were able to find the correct solution, –14. However, 17% of them did not consider a negative solution: they answered 14 (13%) or 0.86/86/6 (2%) or even 0 (2%). A final and somewhat unexpected observation concerns the students who, regardless of their conception of zero, put forward solutions such as ±0.86, ±6, 86 or even –16, as if they were attributing a numerical value to zero such as 1 or even 100. 10% of students who displayed NL conceptions and 6% of those displaying "nothing+" conceptions fell into this category. These students seem to have considered zero as a unit or a hundred and thus to have confused operations under zero with operations in the decimal system (Stacey & Steinle, 2001).

While the results of Table 2 made it possible in particular to examine the type of solution given by the students, those in Table 3 below present the students' success rate with a set of integer additions (Q2) according to their conception of zero (Q1.b).

Conception of zero (Q1.b)	% success with integer additions (Q2)
NL conception	60
"Nothing+" conception	46

Table 3: The relationship between conception of zero (Q1.b) and success with integer additions (Q2)

In table 2, we observe again that students with an NL conception of zero are also more successful at adding integers. The difference is significant (ANOVA: $F=9.92$; $p<0.005$).

CONCLUSION

The "nothing" conception seems to be regarded in scholarship on zero in learning as predominant among students (Levenson et al., 2007; Russell & Chernoff, 2011; Wheeler & Feghali, 1983). Our results reveal a more nuanced reality. Although some of the grade 6 students in our sample – who, it should be remembered, had not yet learned about negative numbers – did demonstrate a "nothing" conception, we saw that an almost equivalent proportion had a conception that related to the number line (NL). The general context of the test about the negative numbers and the MCQ format may have favoured this tendency, but the results also suggest that the "nothing" conception may be less entrenched among students than might appear to be the case. Within the NL conception, the notion of the DNL was especially favoured by the students. While not yet reflecting a unified vision of integers, this conception makes it possible to accept numbers below zero. Our results showed that an NL conception of zero, essentially therefore of the DNL type (in our student sample), was associated with significantly higher success than that associated with a "nothing" conception.

It should be emphasised that a "nothing" conception did not necessarily mean an inability to consider numbers below zero and to perform integer additions; however, the rate of success in performing such calculations was lower. It is also worth noting that, in Q1.a, the correct solution, –14, could also have been arrived at by students who thought that zero "doesn't do anything", and that therefore 0 – 14 = –14, because zero has no effect. A somewhat distinctive conception of zero was apparent in the case of students who gave solutions to 0 – 14 = such as ±0.86, 86, etc., as if they thought zero was equal to a unit or a hundred. Stacey and Steinle (2001) pointed to a similar problem with students who became confused between decimal numbers and negative numbers. According to these authors, these students were confusing the classical model of the number line with the place value columns and treating the spatial arrangement of the usual place value numeration as a kind of "number line" along which the numbers were distributed. We wonder if this confusion could have arisen from students confusing a conception of zero as an empty place indicator with a conception of zero as a number.

Ultimately, it would seem that among students, as in the history of mathematics, working with integers is accompanied by an evolution of the conception of zero towards origin zero (NL), so that zero can be regarded as an integer in its own right. It is therefore vital in school for the extension of natural numbers to integers to go hand in hand with a broadening of the use of zero through a clear explanation of its different functions in different contexts. However, learning about the different meanings of zero does not seem to be part of the curriculum at either primary or secondary level (MENFP, 2011; MENFP, 2008).

REFERENCES

Glaeser, G. (1981). Epistémologie des nombres relatifs. *Recherche en didactique des mathématiques, 2(*37), 303–346.

Levenson, E., Tsamir, P., & Tirosh, D. (2007). Neither even nor odd: Sixth grade students' dilemmas regarding the parity of zero. *The Journal of Mathematical Behavior, 26*(2), 83–95.

MENFP, Ministère de l'Éducation Nationale et de la Formation professionnelle. (2011). *Plan d'études de l'école fondamentale*. Luxembourg.

MENFP, Ministère de l'Éducation Nationale et de la Formation professionnelle. (2008). *Mathématiques : Division inférieure de l'eneignement secondaire. Compétences disciplinaires*. Luxembourg.

Peled, I., Mukhopadhyay, S., & Resnick, L. (1989). Formal and informal sources of mental models for negative numbers. *Proc. of the 13th Conf. of the Int. Group for the Psychology of Mathematics Education* (pp.108–111). PME. Paris.

Russell, G., & Chernoff, E. J. (2011). Seeking more than nothing: Two elementary teachers' conceptions of zero. *The Mathematics Enthusiast, 8*(1), 77–112.

Ruttenberg-Rozen, R. G. (2018). *Growing Prior Knowings of Zero: Growth of Mathematical Understanding of Learners with Difficulties in Mathematics*. Doctoral dissertation. York University,Toronto.

Stacey, K., Helme, S., & Steinle, V. (2001). Confusions between decimals, fractions and negative numbers: A consequence of the mirror as a conceptual metaphor in three different ways. *Proc. of the 25th Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 4, pp. 4–217). PME. Utrecht, The Netherlands.

Toma, M. (2008). A History of Zero. *Journal of Science and Arts, 8*(1), 117–122.

Vlassis, J., & Demonty, I. (2022). The role of algebraic thinking in dealing with negative numbers. *ZDM–Mathematics Education, 54*(6), 1243–1255.

Volken, H. (2000). Le syndrome du précipice. La peur du vide symbolique. Revue européenne des sciences sociales. *European Journal of Social Sciences,* (XXXVIII–119), 23–33.

Wheeler, M. M., & Feghali, I. (1983). Much ado about nothing: Preservice elementary school teachers' concept of zero. *Journal for Research in Mathematics Education,* (3), 147–155.

THE DISCOURSE MAPPING TREE AS A TOOL FOR ANALYZING THE POTENTIAL AND IMPLEMENTATION OF LINEAR ALGEBRA TASKS

Miriam N. Wallach, Einat Heyd-Metzuyanim and Ram Band

Technion - Israel Institute of Technology

We introduce a tool for mapping tasks and their implementation based on the commognitive theory and realization trees. This tool, the Discourse Mapping Tree (DMT), first maps a priori the subdiscourses involved in solving a task, then, a posteriori the discussion of this task in a class. This affords examination of both the mathematical potential of tasks and of how an implementation takes up this potential. We exemplify the DMT on a lesson about linear transformations in a discussion-based linear algebra workshop. The tool highlighted the students' and instructor's role in authoring links between subdiscourses. It also displayed that the instructor was more responsible for meta-level links, while the students made object-level links more readily and easily.

Teaching mathematics includes posing tasks that may offer more or less opportunities for students to engage with mathematical concepts, ideas, and strategies (Sullivan et al., 2015). Examining the potential of such tasks can further our understanding of the disparity between the potential of tasks and their implementation. This is particularly important in the context of cognitively demanding or explorative teaching practices (Smith & Stein, 1998). However, usually the ability to distinguish between the potential of tasks and the take-up in class is difficult. The literature on task design offers a distinction between a priori and a posteriori analysis of tasks (Artigue, 2009). This literature points to the usefulness of detecting through a priori analysis features of tasks that can support certain types of pedagogical goals (e.g. Kieran, 2019).

In line with the task-design literature (e.g. Gravesen et al., 2017), we suggest a tool for mapping the mathematical potential of a task for explorative instruction, made up of two stages: a priori and a posteriori. The a priori stage is independent from the a posteriori analysis thus enabling comparison between the potential and the implementation. Explorative instruction is instruction that affords students maximal opportunities for explorative participation, that is opportunities to author narratives about mathematical objects based on their own reasoning (Weingarden et al., 2019). Encouraging explorative participation, according to Weingarden and colleagues, includes exposing students to multiple realizations of objects and to links between different realizations. These authors suggested a tool, named the Realization Tree Assessment (RTA), to examine the extent to which students were indeed exposed during a whole class discussion to multiple realizations and links between them. However, the RTA did not make a clear distinction between a priori and a posteriori analysis of a task, nor did it precisely define the mathematical potential of a task.

2023. 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. 315-330). PME 46.

In the present work, we build on Weingarden and colleagues' work, and extend it by defining the mathematical potential of a task in correspondence with the socially and historically established mathematical discourse. Inspired by Gee (2015) we distinguish between the mathematical Discourse (with a capital D) - the canonical discourse accepted by the mathematical community, and the classroom discourse (with a small d) - the individualized version of the social Discourse. It is the discourse (with small d) that is seen in discussions around mathematical tasks. Similar to Sfard (2008) we define learning mathematics as becoming a participant in a certain discourse, yet stress that, aligned with Gee, this is a socially and historically established Discourse.

Mathematical Discourses are hierarchical and recursive, where their objects (e.g. $\mathbb{Q}$) build upon previously established objects (e.g. $\mathbb{Z}$) (Sfard, 2008). New mathematical Discourses have historically been created either by several existing Discourses coalescing into one Discourse or by a meta-level Discourse subsuming an older one. Sfard (2008) claims that an individual's adoption of a Discourse, that is learning, often proceeds similarly to how Discourses developed over centuries. Thus, when learners progress from one Discourse to a new subsuming one, the subsuming Discourse includes an isomorphic copy of the old Discourse, as well as new objects and narratives that can only be realized in the new Discourse (Lavie & Sfard, 2019). Adopting new narratives belonging to a familiar Discourse is object-level learning, whereas authoring narratives in the new coalesced Discourse is meta-level learning (Sfard, 2008).

Using this framework, we define the mathematical potential of a task as the potential it affords a student to individualize a Discourse and learning linear algebra as individualizing the Discourse of linear algebra. This includes individualizing all the sub-Discourses of this topic (such as the Discourse of matrices, that of vector spaces, etc.), which we term in this context object-level learning. In addition, it includes linking the realizations from the various sub-Discourses so that the student's individual subsumed discourses coalesce into one discourse, which is considered meta-level learning.

Weingarden and colleagues' (2019) RTA tool constructs a visual representation of realizations of a mathematical object that is at the center of a task, building on Sfard's (2008) definition of a mathematical object being a signifier together with its realization tree. This mapping is constrained, since often multiple mathematical objects are mentioned during a discussion. In the present work, we build on the idea of the RTA to offer a new tool, named the Discourse Mapping Tree (DMT), which enables mapping of a discussion that involves multiple mathematical objects. In addition, we add a clear distinction between the a priori and a posteriori stages of the tool. We conceptualize the building of the a priori DMT as based on an analysis of the mathematical Discourse, whereas the a posteriori stage is based on the classroom discourse that was observed in the lesson.

We apply the Discourse Mapping Tree (DMT) to a single task and to a particular implementation of that task to demonstrate this tool. We ask what the construction of

the DMT highlights about the mathematical potential of the task and about the take up in a whole classroom discussion. In addition, we ask what can the DMT tell us about the potential for and the take up of object-level and meta-level learning?

METHOD

Context, Participants and Data

This study is part of a larger project in which linear algebra workshops were offered to students, in addition to the regular lectures and tutorials and held in parallel to them (Wallach, 2022). The first author led the workshops. Overall, 13 workshops were held, and 7 tasks were designed for them. The tasks were designed by the first and third author, instructors of linear algebra with many years of experience.

The study was conducted at a science and engineering university, where all the students have successfully completed advanced level high-school mathematics courses required for entrance. Students take a linear algebra course, a requirement for most science and engineering students, during their first semester. Participation in the workshops was voluntary and the number of students participating in each workshop varied from 7 to 50. The lesson structure of the workshops was an adaptation of the launch, explore and discuss (LED) structure and Smith and Stein's (1998) suggested practices for orchestrating productive mathematics discussions. Data was collected through several video cameras posed at the board and at student groups. The session described in this paper was one academic hour and 7 students participated.

Analysis

Constructing a DMT includes two parts. First, an a priori analysis of the task, as it is presented to the students, directs the construction (see Fig. 1). This is carried out by experts (mathematicians) who represent the canonical Discourse and is supported by textbooks and curricula. This analysis includes determining the root node, which is the object at the center of the task, listing this object's realizations and finally grouping together realizations of similar type. Each type of realization belongs to a certain Discourse as it has its own keywords, its own narratives, and its own routines of manipulation. Each Discourse (or sub-Discourse) is drawn on a separate branch of the DMT. The a posteriori part of constructing a DMT is based on a video recording of a whole class discussion around the task. In this stage, each narrative authored in a discussion is mapped onto the Discourses identified by the a priori DMT (see Fig. 2). This mapping includes drawing the realizations and links mentioned during the discussion in class. Similar to the construction of RTAs (Weingarden et al., 2019), the components of the DTM demarcate if a student authored it or if the instructor did. Dark boxes and solid lines signal narratives that were authored by students and light boxes and broken lines signal narratives that were authored by the instructor. This process is exemplified in detail in the findings section.

FINDINGS

Mapping the potential of a task

This section describes the process used to construct a DMT and the potential, as revealed by the DMT, for the following task.

> **The linear transformation task** T: $\mathbb{Z}_5^4 \to \mathbb{Z}_5^4$ is a linear transformation such that T(1,2,3,4) = (0,0,0,0). For which values of n ϵ ℕ does there exist such a T so that dim Ker T = n? For those values of n give an example of such a T and find a basis of Ker T.

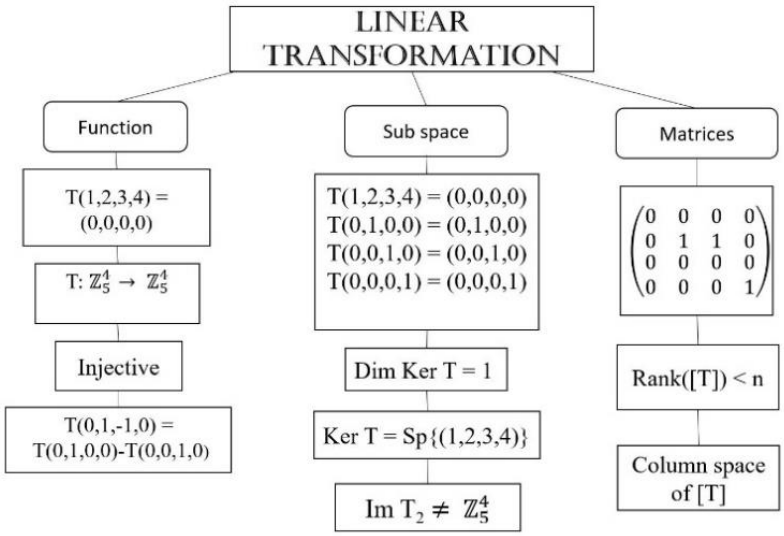
The first step of constructing a DMT (see Fig. 1) is determining the root node. Solving the linear transformation task includes defining a linear transformation with certain properties, thus we determined the root node to be a "linear transformation".

The next step is listing the object's realizations and classifying them into sub-Discourses. We examined the definitions given in textbooks, solutions, and student discourse in workshops, in tutorials, in exams, in homework sets and in questions posed for realizations. These realizations were classified based on the following discourses. A linear transformation is a type of function, thus it can be realized in the Discourse of functions. This includes narratives about the image of vectors, such as T(1,2,3) = (3,3,3), *the image of a vector (x,y,z) is (x+y,x+y,x+y)* and *the linear transformation is injective*. A linear transformation can also be realized using vector spaces. Within this Discourse, a linear transformation can be realized by its definition on any basis. Narratives within the Discourse of vector spaces include *the linear transformation is uniquely determined by defining it on a basis*. Additionally, a linear transformation can also be realized by a matrix representation. The Discourse of matrices includes narratives such as *the linear transformation is invertible since the matrix is invertible.* Linear transformations can also be realized as elements of Hom(V,W). This notion is not included in the curriculum of the linear algebra course examined in this study, thus is not displayed on the DMT. Thus, our a priori DMT analysis revealed that, in this course, the mathematical object linear transformation could be realized in three sub-Discourses – functions, vector spaces, and matrices. Typical realizations were drawn in boxes on the appropriate branches of the DMT (see Fig. 1).

The a priori DMT offers a clear image of the object that can be exposed through the task and its realizations in three different Discourses. Moreover, the DMT demonstrates the opportunity for object level learning and meta-level learning available in this task. Object level opportunities include authoring a realization within a sub-Discourse or saming between two realizations within the same sub-Discourse. For example, the narrative *the kernel is spanned by a single vector and thus the dimension of the kernel is 1* links between two realizations within the sub-Discourse of subspaces. Practicing routines within a sub-Discourse, such as determining the general element of the kernel from a spanning set, is also object-level learning available in this task. Meta-level opportunities include authoring narratives in the coalesced Discourse connecting between two sub-Discourses. For example, the narrative *the image of (1,2,3,4) is (0,0,0,0) so the dimension of the kernel is not zero* connects between a

realization in the functions sub-Discourse (T(1,2,3,4) = (0,0,0,0)) and a realization in the vector space sub-Discourse (dim Ker T ≠ 0). Such a narrative, if authored during the discussion, provides an instance of coalesced discourse which is part of the meta-level learning sought through this task.

Figure 7 - DMT for linear transformation task



Mapping the implementation of a task

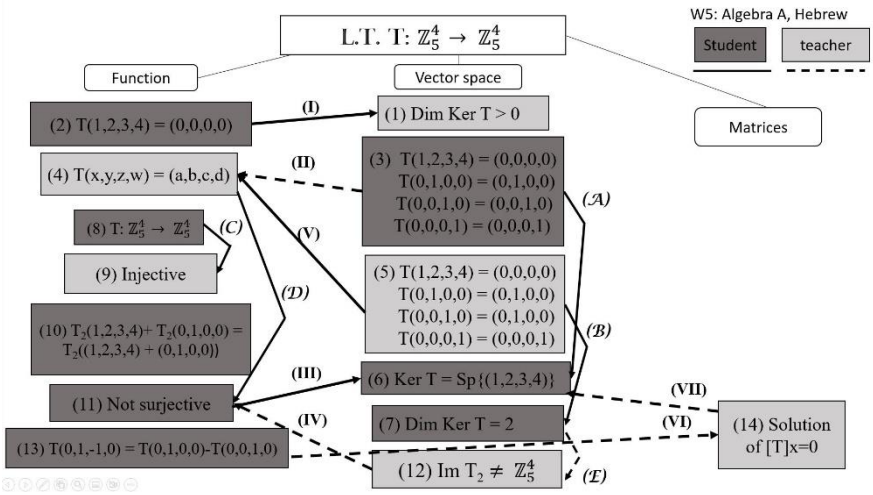
We now exemplify how an implementation of the task was mapped onto the DMT. The workshop on which we focus started with a short reminder of the basic theorems and definitions pertaining to linear transformations. Some of these were in the sub-Discourse of functions and some were in the sub-Discourse of vector spaces. The students were familiar with these narratives from the previous lectures and tutorials. After the launch of the task, the students worked on the task in pairs for 15 minutes. This was followed by a whole class discussion that was 21 minutes long. We use the DMT to map the whole class discussion. There were seven students in the classroom, and they all participated in the discussion. Some talked from their seats, and some came to the board to write out examples or to point to examples already written.

The mapping of the implementation commences by deriving the node and the sub-Discourses from the a priori DMT, described in the previous section. The realizations that were used to determine the sub-Discourses for the DMT are removed. They indicate hypothetical narratives, that may or may not be authored in the class and are only used as examples to map the sub-Discourses and potential links between them. The a posteriori DMT analyzes the discourse in the discussion and maps the narratives authored in class onto the sub-Discourses identified and the connections made between these narratives. Figure 2 presents the DMT mapping of the whole class discussion.

The a posteriori DMT in Figure 2 shows that while there were three available sub-Discourses, the discussion included mostly narratives that belonged to only two of them. The sub-Discourse of matrices, mentioned only briefly by the instructor, was not used by the students at all. In the other two sub-Discourses, the realizations and links

were authored both by the instructor and by the students in a mostly balanced manner. Yet there is a noticeable difference between these two sub-Discourses. In the functions sub-Discourse, five out of the seven realizations mentioned were authored by the students, signalling they favored this sub-Discourse, whereas in the subspace sub-Discourse the division is more equal.

Figure 8 - DMT for whole class discussion



We now exemplify the take up of object-level and meta-level learning that occurred during the whole class discussion. During the discussion a student said, “We can define the linear transformation by its behavior on the basis” and wrote this on the board (see Fig. 2, box 3). The instructor agreed and added that the final answer would need to be given for a general vector (4) similar to a function. The instructor outlined how to do so and connected between the realizations (II). The discussion then turned to the kernel of this transformation and a student stated that the kernel is the span of (1,2,3,4) (6, $\mathcal{A}$). The discussion that next ensued about the linear transformation on a general vector (4) within the function subdiscourse sparked a student’s question does such a definition define a function. The student’s explanation of his question and the instructor’s statements elicited from other students that the transformation is not surjective ($\mathcal{D}$,11). The instructor then asked for a connection to the kernel (6), written on the board previously and students authored this (III).

In the excerpt described there was both object-level learning and meta-level learning. The students authored object-level links ($\mathcal{A}$,$\mathcal{D}$) within each sub-Discourse. The instructor, in contrast, stressed the meta-level links (II,III) and put less emphasis on the object-level narratives. More generally, the DMT displays that the students authored more object-level narratives during this discussion. The instructor moved the discussion to include meta-level narratives, as seen by the four meta-level inks authored by the instructor (II, IV, VI, VII), as opposed to a single object-level link ($\mathcal{E}$). Additionally, the links authored by the instructor show that when the discussion was in the functions sub-Discourse, the instructor pushed it to the other ones (II, IV, VI). The potential for object-level learning was taken up by the students authoring

narratives mostly within the function sub-Discourse. The potential for meta-level learning was offered mainly by the instructor authoring or supporting the students to author meta-level narratives between sub-Discourses.

DISCUSSION

This study examined what the construction of the DMT highlights about the mathematical potential of the task and about the take up of this potential. The mathematical potential of the task was analyzed based on a priori analysis of mathematical Discourses involved in the task, and an implementation of the task was examined a posteriori based on the discourse that occurred during the classroom discussion. We found that constructing the DMT a priori emphasized the opportunities for both object-level learning and meta-level learning afforded by the task. Additionally, mapping the task afforded a structured look at the various mathematical sub-Discourses involved in solving the task. The a posteriori DMT displayed an image of the whole class discussion that showed how the potential for object-level learning and meta-level learning offered by the task were taken up in an implementation.

The construction of these DMTs highlighted the difference between the expectations from the task and the implementation of the task in terms of the mathematical potential. For example, while the DMT of the task showed that the potential for involving three sub-Discourses in the discussion was available, the implementation only focused on two sub-Discourses. The highlighting of this neglect shows the affordances of a tool which can determine the mathematical potential of a task and the take up of this. The DMT tool's examination of the mathematical potential of a task, independent of the context in which it will be used, allows analysis of tasks before they are implemented in a classroom setting and supports choosing appropriate tasks for use.

The DMT also helps to differentiate between object-level learning and meta-level learning. In the workshop discussion we analyzed, the a posteriori DMT showed that the instructor was more responsible for authoring meta-level links. The students easily and more readily authored object-level links, with little or no support from the instructor. This aligns with Nachlieli and Elbaum-Cohen's (2021) suggestion that student-centered instruction might support meta-level learning when strongly guided by an instructor who can explicate the rules of the new discourse and stress the limitations of the old, familiar discourse.

In this paper, we applied the DMT to one lesson, mainly to illustrate its construction. Applying the DMT to multiple lessons affords comparison of various aspects of the whole class discussion, similarly to was been achieved with the RTA (Weingarden et al., 2019). In the larger study, constructing DMTs for multiple workshops strengthened the impressions from the present analysis that the links between sub-Discourses were dependent on the students' familiarity with the narratives within the sub-Discourses and that there was usually a dominant Discourse which is either more familiar to the students or which includes familiar procedures (e.g. the function subdiscourse in this study) (Wallach, 2022). Constructing DMTs can thus be productive in examining

various aspects of whole class discussions, especially in explorative and learner-centered forms of instruction.

REFERENCES

Artigue, M. (2009). Didactical design in mathematics education. *Nordic Research in Mathematics Education, Brill,* 5–16.

Gee, J. P. (2015). Discourse, small d, big D. *The International Encyclopedia of Language and Social Interaction, 3,* 1–5.

Gravesen, K. F., Grønbæk, N., & Winsløw, C. (2017). Task Design for Students' Work with Basic Theory in Analysis: the Cases of Multidimensional Differentiability and Curve Integrals. *International Journal of Research in Undergraduate Mathematics Education, 3*(1), 9–33.

Kieran, C. (2019). Task Design Frameworks in Mathematics Education Research: An Example of a Domain-Specific Frame for Algebra Learning with Technological Tools. In G. Kaiser & N. Presmeg (Eds.), *Compendium for Early Career Research in Mathematics Education, ICME 13 Monographs* (pp. 265–287).

Lavie, I., & Sfard, A. (2019). How children individualize numerical routines: Elements of a discursive theory in making. *Journal of the Learning Sciences, 28*(4–5), 419–461.

Nachlieli, T., & Elbaum-Cohen, A. (2021). Teaching practices aimed at promoting meta-level learning: The case of complex numbers. *Journal of Mathematical Behavior, 62,*100872.

Sfard, A. (2008). *Thinking as Communicating: Human development, the growth of discourses, and mathematizing*. Cambridge University Press.

Smith, M. S., & Stein, M. K. (1998). Reflections on practice: Selecting and creating mathematical tasks: From research to practice. *Mathematics Teaching in the Middle School, 3*(5), 344–350.

Sullivan, P., Knott, L., & Yang, Y. (2015). The relationships between task design, anticipated pedagogies, and student learning. In A. Watson & M. Ohtani (Eds.), *Task Design in Mathematics Education, New ICMI Study Series* (pp. 83–114).

Wallach, M. N. (2022). *Encouraging Explorative Participation in Linear Algebra through Discourse-Rich Instruction - the Importance of Object-level and Meta-level Learning*. Technion- Israel Institute of Technology.

Weingarden, M., Heyd-Metzuyanim, E., & Nachlieli, T. (2019). The realization tree assessment tool – examining explorative participation in mathematics lessons. *Journal of Mathematical Behavior, 56*100717.

ANSWER PATTERNS OF JAPANESE SECONDARY SCHOOL STUDENTS IN TIMSS 2015 MATHEMATICS SURVEY

Koji Watanabe[1]

[1]Miyazaki International College

This study used secondary analysis of the Trends in International Mathematics and Science Study (TIMSS) 2015 eighth-grade mathematics survey data to clarify answer patterns of Japanese secondary school students. While these patterns are established at the primary school level, they have yet to be set for Japanese secondary school students. In this study, to identify students' answer patterns, we performed an international comparative analysis of 15 countries and areas. The results showed that Japanese secondary school students had a unique answer pattern in comparison to these 15 countries and areas. In particular, the 'Number' items, especially those learnt at the primary school level, were found to be more difficult for Japanese students than for students from other countries and areas.

INTRODUCTION

The Programme for International Student Assessment (PISA) and Trends in International Mathematics and Science Study (TIMSS) are influential international educational assessments in mathematics education in Japan (e.g., Nakayasu, 2016; Volante, 2015). They began to accumulate data in 2000, providing resources for secondary data analysis, to gain new insights into Japanese mathematics education.

Previous research on Japanese students' mathematical achievements such as Suzukawa et al. (2008) indicated that Japanese students had unique answer patterns; they did especially well at solving questions in the 'educational' context of the PISA framework, as compared to data from 13 countries and areas (i.e., Australia, Canada, Finland, France, Germany, Hong Kong, Ireland, Italy, Japan, Korea, the Netherlands, New Zealand, and the United States). Watanabe (2019, 2020) confirmed the unique overall answer patterns provided by Japanese students compared to the same 13 countries and areas. Further, this research identified a partial change: item difficulty in the mathematical content 'uncertainty and data' decreased between the PISA 2003 and the 2015 study.

Meanwhile, in the secondary analysis of TIMSS mathematics survey data, considering the answer pattern, Watanabe and Watanabe (2021) identified the answer patterns of Japanese fourth-grade students in TIMSS 2015, for the same 13 countries and areas, to show that Japanese primary school students had a unique answer pattern; specifically, calculation items were found to be easier for Japanese students than for students from the other countries and areas. Although Japan participates in the G8 survey, there is no mention of the eighth-grade students' answer patterns. Thus, this study aimed to reveal the answer patterns of Japanese secondary school student using TIMSS eighth-grade data.

2023. 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. 323-330). PME 46.

The percentage of correct answers is often used to identify answer patterns while focusing on item difficulty. The National Institute for Educational Policy Research (NIER, 2017) shows the percentage of correct answers given for each item in the mathematics survey. However, simply comparing the percentage of correct answers in each country with that in Japan is not sufficient to determine the answer pattern as the percentage of correct answers is not only an indicator of the item difficulty; it also depends on the group of examinees. Thus, separating the item characteristics from the examinee characteristics is necessary to focus on item difficulty and enable the identification of answer patterns. Item response theory (IRT) is known as the test theory that makes distinguishing characteristics possible (Henard, 2000). The method of analysis using IRT conducted by Suzukawa et al. (2008) and Watanabe and Watanabe (2021) were referred to in this study.

METHODS

As Watanabe and Watanabe (2021) analysed TIMSS 2015 data, this study did the same. However, some of the 13 countries targeted by Watanabe and Watanabe (2021) did not participate in the eighth-grade mathematical survey. Instead, 15 countries and areas (i.e., Australia, Canada, Hong Kong, Hungary, Ireland, Italy, Japan, Korea, New Zealand, Russia, Singapore, Sweden, Taiwan, the United Kingdom, and the United States) participated in the eighth-grade mathematical survey and reported their results in TIMSS 2015, published by NIER (2017). The present analysis targeted 91,256 eighth-grade students from these 15 countries and areas.

In TIMSS 2015, 14 different booklets were prepared, and 212 items were given to measure achievements in mathematics learned in school. Three items (item codes: M062345B, M062342, and M062342) were not applicable and were excluded from scaling at the international level. Two items (item codes: M062237 and M052090) were not applicable and excluded from scaling at the national level in Sweden (Martin et al., 2016), to ensure the precision of comparison. In total, 207 items were included, out of which 13 had partial credit; answers with partial credit were treated as incorrect to avoid complicating the data analysis. A binary dataset (1 for a correct answer and 0 for an incorrect answer, non-response, or missing answer), was built for this study. Incidentally, items not included in the booklets given to the students were regarded as not available (NA) during the statistical analysis that was performed using the statistical data analysis software R version 4.0.5.

This analysis applied the Rasch model of IRT, expressed as follows:

$$p_i(\theta) = \frac{1}{1+\exp\{-1.702(\theta - b_i)\}}$$

where θ is the latent trait of ability, p_i denotes the probability of whether an answer to the item i is correct, and b_i denotes the difficulty parameter of the item i. The item difficulty parameters b_i (i =1, 2, …, 207) of 207 items were estimated for each country, using the Rasch model and compared by equating it to the scale of Japan with the mean-sigma method. More specifically, let the item difficulties of the item i for Japan and

country k (k =1, 2, …, 15) be b_{iJPN} and b_{ik}, respectively, and let the mean values of item difficulties be $\overline{b_{JPN}} = \frac{1}{207}\sum_{i=1}^{207} b_{iJPN}$ and $\overline{b_k} = \frac{1}{207}\sum_{i=1}^{207} b_{ik}$, respectively. Then, the item difficulties, equated to the Japanese scale, are defined as $b_{ik}^* = b_{ik} + (\overline{b_{JPN}} - \overline{b_k})$. As $\overline{b_k^*} = \overline{b_{JPN}}$ can be obtained, the mean value of the equated item difficulties is combined into $\overline{b_{JPN}}$. Let the mean value of the item i in 15 countries be $\overline{b_i^*} = \frac{1}{15}\sum_{k=1}^{15} b_{ik}^*$. Assuming a country with the difficulty $\overline{b_i^*}$ for each item, it is possible to set up a country with an average pattern of item difficulty across 15 countries (hereinafter referred to as 'average country'). Given the difference in item difficulty between each of the 15 countries and the average country, $d_{ik} = b_{ik}^* - \overline{b_i^*}$, we obtain $\overline{d_k} = \overline{d_i} = 0$. In this manner, d_{ik} is obtained as a standardised item difficulty for each country and item and is used as an indicator of the uniqueness of an answer pattern in comparison to the 15 countries. This analysis focuses on d_{ik}, to detect the answer patterns of 15 countries. The analysis mainly used the packages 'ltm' and 'plink' in the statistical data analysis software R version 4.0.5 (Rizopoulos, 2018; Weeks, 2017).

RESULTS

Overall Features of Answer Pattern in 15 Countries and Areas

Let the standard deviation of d_{ik} be $s_k = \sqrt{\frac{1}{207}\sum_{i=1}^{207} d_{ik}}$, where s_k is an indicator of the difference in the overall level of item difficulty for the 15 countries and areas. A larger s_k indicates a more divergent answer pattern for the corresponding country. Table 1 lists the s_k values obtained in this study.

Table 1 indicates that Japan has a high s_k value, which is less than that of Taiwan and Korea. Additionally, we found that East Asian countries, such as Taiwan, Korea, Japan, Singapore, and Russia, have divergent answer patterns. Thus, Japan is inferred to have a unique answer pattern among the 15 countries and areas in eighth grade TIMSS 2015 mathematics surveys.

Table 1: Values of s_k

Country	s_k	Country	s_k	Country	s_k
TWN	0.415	SWE	0.359	ENG	0.287
KOR	0.413	ITA	0.318	NZL	0.265
JPN	0.409	HKG	0.310	HUN	0.258
RUS	0.372	IRL	0.299	AUS	0.247
SGP	0.362	USA	0.290	CAN	0.228

Relationship Between Item Difficulty and Item Content

The items in the TIMSS 2015 are characterised by two aspects: cognitive domains and content domains. The characteristics of the Japanese answer patterns were examined by focusing on these two aspects. The cognitive domains include three content areas: Applying (92 items), Knowing (69 items), and Reasoning (46 items). The content domains contain four types of content: Algebra (59 items), Data and Chance (41 items), Geometry (43 items), and Number (64 items). The details of these definitions are provided in Mullis et al. (2013, pp. 19-27). The distribution of d_{ik} for each of these two aspects was checked using a boxplot, as shown in Figures 1 and 2.

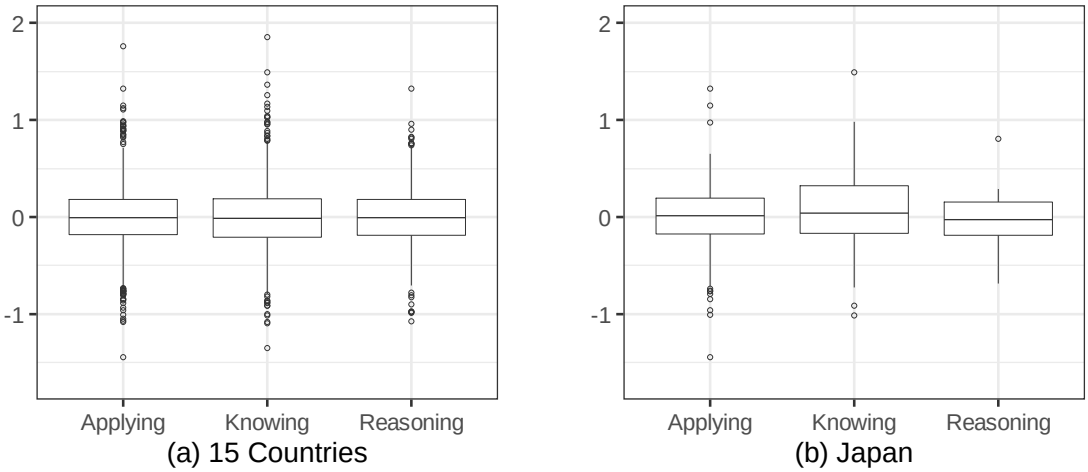


Figure 1: Boxplots for the cognitive domains

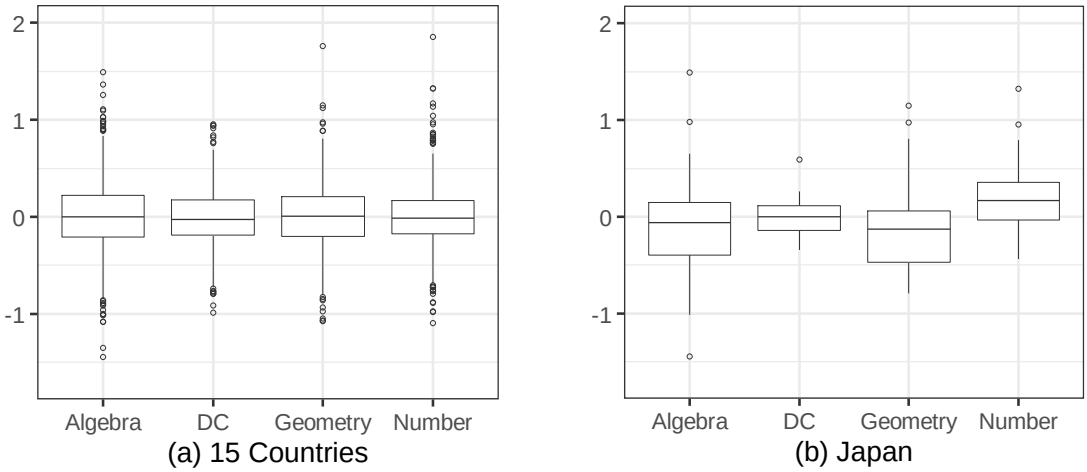


Figure 2: Boxplots for the content domains (DC = Data and Chance)

Figures 1 and 2 show boxplots of the cognitive and content domains, respectively. The boxplots in Figures 1(a) and 2(a) depict the distribution for all 15 countries, while those in Figures 1(b) and 2(b) depict the distribution for Japan. The median values for the cognitive and content domains were close to 0.0 in both Figures 1(a) and 2(a). This feature can only be observed for Japan in Figure 1(b). The box showing the cognitive domain content area, listed as ‘Knowing’ in Figure 1(b), is slightly larger than it is in

the 15 countries and areas. However, there is no significant difference in the distribution of d_{ik} across these countries, not even in Japan, as shown in Figure 1(b).

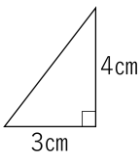
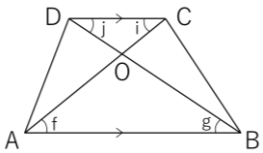
The differences in the distribution of d_{ik} are shown in Figure 2. The distribution of d_{ik} for the content domain listed as 'Number' for Japan in Figure 2(b) is higher than it is for the 15 countries and the areas in Figure 2(a). This implies that these items are relatively difficult for Japanese students. A characteristic of Japanese answer patterns was found in the content domain 'Number' of the TIMSS 2015 framework.

Analysis Focused on Released Items

In total, 93 out of 207 items were released to the public. Based on these, we examined the answer patterns of Japan. As the value of $s_{JPN}$was 0.409 (Table 1), the released items with a larger value of s_{JPN} values are given in Table 2.

Four of the 10 items were not yet been learned by Japanese students when the TIMSS 2015 was conducted (NIER, 2017). Predictably, these four items would be highly difficult for Japanese students.

Japanese students have already learned the other six items: specifically, five under the content domain 'Number', and one under 'Algebra', at school. The five items M042019, M042060, M052209, M042183, and M042302A refer to answering addition with an approximate number, providing a fraction that represents a percentage, expressing a fraction as a decimal, finding the multiplier, and performing the appropriate calculation, respectively. Japanese students learn about these in primary schools. Thus, the items studied in primary schools were found to be relatively difficult for Japanese students.

Item code	Domain	Summary of item content	d_{iJPN}
M052103*	Algebra	Finding the value of y: $y = \sqrt{x-9}$, when $x = 25$.	1.495
M052042*	Geometry	 Finding the length of the remaining side of the triangle.	0.973
M042264*	Geometry	 Triangle A and B are homologous. Finding a pair of equal angles.	0.807
M042019	Number	Finding the answer for 103+289+475+310+519 to the approximate number to the hundredth place.	0.791
M042060	Number	Finding the fraction representing a 20% discount.	0.584
M052064	Algebra	Finding the value of $\frac{a^2}{2} - 6a + 36$, when $a = 3$.	0.583
M052209	Number	Finding the closest decimal number in size to $\frac{3}{4}$.	0.526

M042183	Number	Finding the number which, when dividing 202.6 by that number, gives the answer 2.026.	0.476
M042302A	Number	Plan of company X: • Basic monthly fee of 250 zeds • Calls cost 3 zeds per minute • 2 zeds per email Plan of company Y: • 2500 zeds for the mobile phone • Basic monthly fee of 50 zeds • Calls cost 2 zeds per minute • 1 zeds per email Finding the cost per year for Company X and Company Y, if you do not make any calls and do not send any emails.	0.475
M042248*	Algebra	Finding the value of n, if $\frac{8}{12} = \frac{24}{2n}$.	0.454

* Not taught at school at the time TIMSS 2015 was conducted.

Table 2: The items relatively difficult for Japanese students compared with 15 other countries and areas (Source: NIER, 2017)

DISCUSSION

We conducted a secondary data analysis of the TIMSS 2015 eighth-grade mathematics survey to identify the characteristics of the answer patterns of Japanese secondary school students. Standardised item difficulties d_{ik} were calculated, based on which an international comparative analysis was conducted.

In summary, the following two points were identified as the main findings. First, the overall answer pattern at the eighth-grade secondary school level in Japan was found to be unique in comparison to 15 countries and areas, after Taiwan and Korea (Table 1). Second, using the framework of the TIMSS 2015, we examined these peculiarities and found that the content domain ‘Number’ was relatively difficult for Japanese eighth-grade students (Figures 1 and 2). More specifically, we examined 93 released items and identified ten items that Japanese students found difficult to complete. Besides the four items that the students had not learned at school, the five items of the content domain listed as ‘Number’ were found to be difficult for Japanese students (Table 2).

Watanabe and Watanabe (2021) reported that, in terms of the answer patterns of fourth-grade Japanese primary school students in the TIMSS 2015, items such as addition, subtraction, and division of integers and decimals, and solving simple equations (for example, 0.36+0.77 and 1362÷32), which are taught in primary school, are less difficult. Suzukawa et al. (2008) reported that, in the framework of the PISA 2003, questions in educational contexts, defined as typical items learned in school, are easier for Japanese students. Meanwhile, the three least difficult items in this analysis were studied at the secondary school level (see Table 3).

Based on the previous studies and Table 3, it can be concluded that the answer patterns of Japanese eighth-grade secondary students suggest they can easily solve the content of items learnt in secondary school, but not those learnt in primary school or those not

yet learnt. It is especially necessary to ensure that the content domain listed as ‘Number’, taught in primary school, is retained through secondary school.

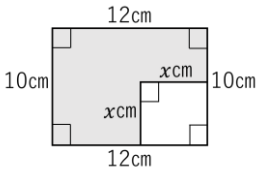
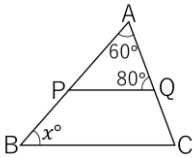
Item code	Domain	Summary of item content	d_{iJPN}
M052126	Algebra	 Finding an expression in terms of x for the area of the shaded portion of the figure.	-0.960
M052131	Algebra	Finding the right process to solve the equation $4x - 3 = 2x - 7$.	-0.799
M052417	Geometry	 The line PQ and BC are parallel. Finding the angle x.	-0.743

Table 3: The items easier for Japanese students compared to 15 countries (Source: NIER, 2017)

CONCLUSION

Through this study, we established specific answer patterns of Japanese secondary school students using IRT. For example, the items under the content domain ‘Number’, are relatively difficult; some items taught at the primary school level are difficult for Japanese students at the secondary level. However, they demonstrated decent knowledge of the items learned at secondary school. Nevertheless, not having focused on items that are easy for Japanese students, we believe that efforts are required to capture the characteristics of Japanese answer patterns in greater detail; for example, it is necessary to examine the relationship between item difficulty and item content. In addition, Watanabe (2020), Watanabe and Watanabe (2021), and this study have revealed the Japanese answer patterns in the PISA and TIMSS, especially the PISA 2015 mathematical literacy survey and the TIMSS 2015 fourth- and eighth-grade mathematics surveys. A synthesis of these results is warranted for future work.

Moreover, the TIMSS is conducted every four years, and its accumulating data makes it possible to describe changes over time. For example, it would be interesting to identify whether the specific challenge of ‘Number’ items also appeared in TIMSS 2011 or TIMSS 2007; that is, is this a stable pattern or just an incidental result for TIMSS 2015? In addition, the spread of COVID-19 has had a significant impact on schooling and is presumed to affect students’ actual achievements. Visualizing its impact is also important for future research. While this study provides information on students’ achievements before the spread of COVID-19, future research may include clarification of changes over time among the TIMSS 2015, 2019, and 2023, respectively, the last of which is scheduled to be implemented in 2023.

ACKNOWLEDGEMENT

This work was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI Grant Number 20H01689.

REFERENCES

Henard, D. H. (2000). Item response theory. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding MORE multivariate statistics*, pp. 67–97. American Psychological Association.

Martin, M. O., Mullis, I. V. S., & Hooper, M. (2016). *Methods and Procedures in TIMSS 2015*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I. V. S., & Martin, M. O. (2013). *TIMSS 2015 Assessment Frameworks*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Nakayasu, C. (2016). School Curriculum in Japan. *The Curriculum Journal*, 27(1), pp. 134-150.

NIER (National Institute for Educational Policy Research) (2017). *TIMSS2015 sansuu suugaku kyouiku rika kyouiku no kokusai hikaku: kokusai suugaku rika kyouiku doukou chousa no 2015nenn chousa houkokusho [TIMSS 2015 International Comparison of Mathematics and Science Education]*. Tokyo: Akashishobo (in Japanese).

Rizopoulos, D. (2018). Package 'ltm': Latent Trait Models Under IRT. Retrieved from https://cran.r-project.org/web/packages/ltm/ltm.pdf.

Suzukawa, Y., Toyoda, H., & Kawahashi, I. (2008). Do Students in Japan Learn How to Use Mathematics in Real-Life Situations? An Analysis of Items in the 2003 Programme for International Student Assessment (PISA 2003). *The Japanese Journal of Educational Psychology*, 56(2), pp. 206-217. (in Japanese).

Volante, L. (2015). The Impact of PISA on Education Governance: Some Insights from Highly Reactive Policy Contexts. *International Studies in Educational Administration*, 43(2), pp. 103-117.

Watanabe, K. (2019). A Study on the Secular Changes of Japanese Students' Mathematical Literacy: Focusing on the Pattern of Item Difficulty Between PISA 2003 and PISA 2012. *Journal of Japan Academic Society of Mathematics Education* (*JASME*). *Research in Mathematics Education*, 25(1), pp. 1-13. (in Japanese).

Watanabe, K. (2020). A Study on the Changes of Japanese Students' PISA Mathematical Literacy: Focusing on the Content Category of 'Uncertainty and Data' on PISA 2003, PISA 2012 and PISA 2015. *Journal of Japan Academic Society of Mathematics Education* (*JASME*). *Research in Mathematics Education*, 26(1), pp. 1-12. (in Japanese).

Watanabe, K., & Watanabe, Y. (2021). Answer Patterns of Japanese Primary School Students in TIMSS 2015 Mathematics Survey, *Proceeding of the 44th Conference of the International Group for the Psychology of Mathematics Education* Vol. 4, pp.238-246. Khon Kaen, Thailand: PME.

Weeks, J. P. (2017). Package 'plink': IRT Separate Calibration Linking Methods. Retrieved from https://cran.r-project.org/web/packages/plink/plink.pdf.

REPRESENTING 'TALL AND SHORT' IN DRAWINGS – PRE-SCHOOL TO YEAR 2

Jennifer Way

University of Sydney, Australia

The study reported in this paper contributes to the exploration of the development of children's mathematical drawing. 109 Preschool-to-Year 2 Australian children (4 to 8 years) produced drawings of 'something tall and something short'. Open-ended analysis of the forms and structures of the drawings revealed five categories ranging from scribble to the base-line comparison of two objects. The variety in drawing forms and the scattering of ages across the five categories suggests that educators should be more aware of children's drawing development in association with their skills in creating mathematical representations.

INTRODUCTION AND BACKGROUND LITERATURE

Drawing is a fundamental form of representation in mathematics, and from the commencement of formal schooling there is an expectation that children will increasingly make use of conventional drawing techniques and diagrams. However, there is little guidance available to teachers about what representational forms to expect from children and how to support the early development of mathematical representation, particularly drawing (Bobis & Way, 2018). As emphasised by Ginsburg, Lee and Boyd (2008), children need to be supported in 'mathematising' their self-created representations. A currently under-researched aspect of mathematical development is how young children transition from natural drawing, to drawing as a mathematical representation tool (Way, 2018).

The emergence of drawing as a representational form

Young children's drawing develops in its purpose and form over several years, moving in stages from playful scribble and exploration of forms, to pictorial and iconic representations of visualizations and real-world objects. Scribble is therefore a natural part of drawing development, that gradually takes on more controlled forms such as wavy lines, circles and dashes; followed by composite figures including a human ideogram around 3-4 years (Carruthers & Worthington, 2005; Machón, 2013). Typically, children do not begin deliberately representing external objects (symbols and pictures) until 4-5 years. Gradually the emergence of schemas (pictorial structures) such as baselines, proportionality and the exploration of perspective can be seen in children's drawings in the 4 to 7 years range (Machón, 2013). At this stage in drawing development, children may have the potential to apply 'natural' drawing to the representation of mathematical concepts and processes, however, it should not be assumed that all children can create mathematical drawings without explicit support (Way, 2018). Indeed, the role of interactions with adults to support semiotic activity (reflecting on the relationship between the sign and its meaning) has been highlighted

2023. 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. 331-338). PME 46.

by several researchers (e.g., Papandreou, 2014; van Oers, 1997). With a focus on explicitly supporting schematising activities with 5/6-year-olds, Poland and van Oers (2007) established significantly better schematising and mathematics skills a year later, compared with children who had not received similar support.

Young children's drawings of length comparison

Young children develop a range of contextualised measurement concepts through their informal experiences, prior to engaging in more formal instruction at the commencement of school (Chigeza & Sorin, 2016; MacDonald & Lowrie, 2011). In a rare study of young children's representation of mathematical ideas through drawing, MacDonald (2010) examined the drawings of 'something tall and something short', produced by 83 children who had just commenced formal schooling at two Australian schools (around the age of 5 years). The children were also asked to provide a verbal description of the drawings, facilitating richer data and more accurate interpretations of the children's contextualised drawings. The majority of children (80 to 90%) produced representations that; a) focused on the specific attribute of length, b) made direct comparison of the heights of two objects, and, c) used appropriate measurement language, often including comparative terms such as taller and shorter (MacDonald, 2010). Similarly, Chigeza and Sorin (2016) found that 4/5-year-olds in Australia and Canada were able to focus on a selected attribute of objects, such as length, and demonstrate comparison through drawings and words.

Visualization and concept images

Visualization has become widely accepted as a vital component of mathematics education (Acarvi, 2003). Drawing is a form of external visual representation related to internal mental imagery and other thinking processes (Papandreou, 2014). In the context of exploring young children's emerging, and highly contextualised mathematical understandings, Tall and Vinner's (1981) notion of 'concept images' provides an apt theoretical perspective for considering the visual representations (drawings) of particular concepts - in this case, height (specifically, tall and short). The term 'concept image' refers to,

> "… the total cognitive structure that is associated with the concept, which includes all the mental pictures and associated properties and processes. It is built up over the years through experiences of all kinds, changing as the individual meets new stimuli and matures" (Tall & Vinner, 1981, p. 152).

Different stimuli at a particular time may activate only part of the concept image to produce an 'evoked concept image'. Tall and Vinner (1981) use this to help explain students' conflicting responses to different tasks in relation to the same concept. An implication for this paper is that caution is needed when interpreting the mathematical aspects of a visual representation created by a child. In addition to accounting for drawing development, it might be difficult to distinguish between an incomplete 'concept image' (the source) and the 'evoked concept image' (the selected part).

Focus of the Study

The purpose of this paper is to reveal the range of drawings produced by children in response to a simple verbal prompt regarding relative height. Although understanding of 'tall and short' can, to some extent, be inferred from the drawings, the emphasis is on the nature of the drawings themselves rather than the assessment of mathematical concepts. Therefore, this paper is focused by the question: What types of drawings do 4- to 8-year-old children create to represent 'tall and short'?

METHOD

The study reported in this paper is part of a larger ongoing research project, *Emerging Mathematical Drawing* that explores the development of mathematical drawing across preschool and primary years, with particular interest in the transition from naturalistic drawing to mathematical diagrams. The site for this 2022 study was a state primary school with an attached preschool in a low socio-economic metropolitan area of a major city in Australia. Data collection took place in the 4th month of the school year. All the children attending the participating classes on the day were invited to complete the drawings (See Table 1), and the drawings from children whose parents provided written consent were used as data.

Table 1. Participant information. Note: 'Foundation' is the generic name used by the Australian curriculum for the first year of mandatory school attendance.

School level	Approximate age range	Number of participants
Pre-school	4-5 years	10 from 1 class
Foundation	5-6 years	32 from 3 classes
Year 1	6-7 years	36 from 2 classes
Year 2	7-8 years	31 from 2 classes

The teacher was present, but the instructions were delivered by the researcher as a stand-alone activity – that is, the drawings were not contextualised in a play experience of imbedded in a lesson on measurement.

> Verbal instructions: "Think about things that are tall (raise hand over head height) and things that are short (lower your hand below waist level). Draw something tall and something short".

An open-ended inductive approach was used to gradually sort the 109 drawings into groups as obvious similarities began to emerge on examination, such as the number of figures, recognisable objects, scribble etc. These tentative groupings were then examined more closely for the use of schemas, that is, pictorial structures such as proportion and the use of a baseline. These schemas overlap with the mathematical characteristics of depicting of tall/short and the use of a baseline for clear height comparison. At this point a second researcher with prior experience in analysing children's mathematical representations reviewed and discussed every drawing, which resulted in the refinement and consolidation of five groupings.

FINDINGS

Category A: Incoherent

The 21 drawings (19.3%) in this category offered no discernible representation of the concepts of tall and short, including one ‘no response’. Some of the drawings were scribble, in the developmental sense, either freeform scribble or with some form and structure (see Figure 1). It is possible that some of the drawings were informed by visualised ‘tallness; or ‘shortness’ but such a representation could not be inferred.

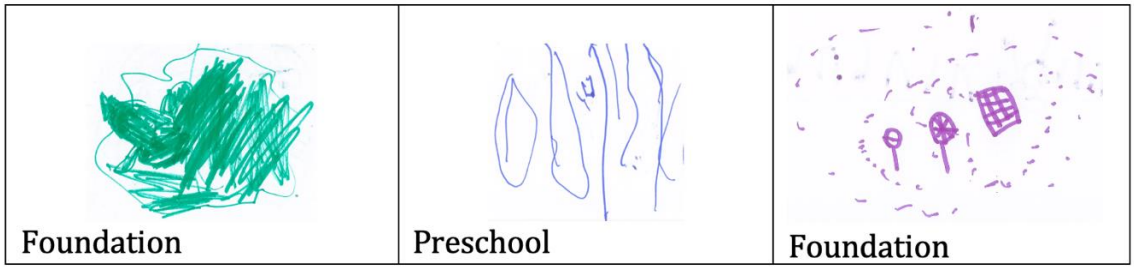


Figure 1. Examples of Category A: Incoherent

Category B: One object

Only eight drawings (7.3%) were of a single object that could be interpreted as a ‘known’ tall object (e.g., a tree, giraffe, adult) or a ‘known’ small object (e.g., and insect, a cat). In these drawings, no comparison of height was depicted but it was assumed that the tallness or shortness of the object was relative to the child’s own height (See Figure 2).

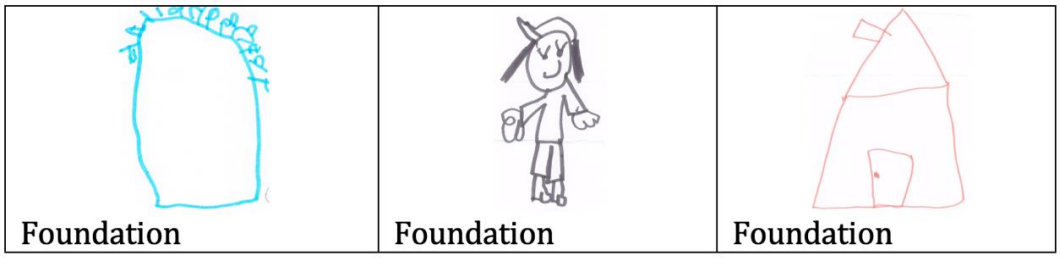


Figure 2. Examples of Category B: One object

Category C: Similar size

The seven drawings (6.4%) in this category contained two or multiple figures of the same or similar height or length and were spread across the four year-levels. These drawings suggested awareness of the height/length concept and perhaps of comparison, but the drawing did not clearly depict the difference in height expected when illustrating tallness and shortness. Some drawings were of lines or geometric shapes rather than physical objects (see Figure 3). Some included a baseline.

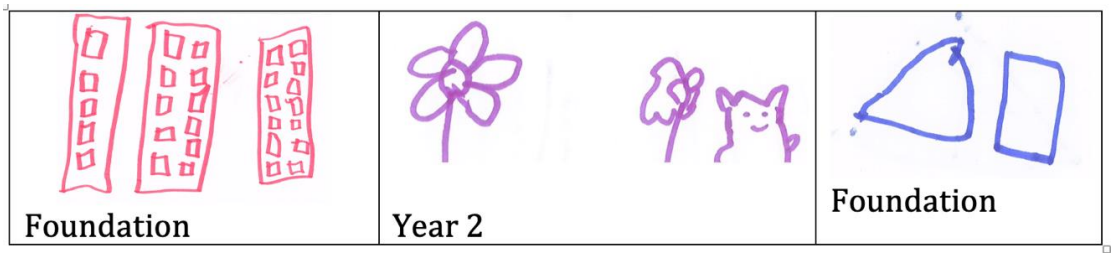


Figure 3. Examples of Category C: Similar size

Category D: Figures not on a baseline

The 41 drawings (37.6%) in this category contained representation of two objects (or multiple figures) of clearly different heights/lengths but did not include clear base line for comparison. Some drawings were lines or geometric shapes, and a few were presented horizontally rather than vertically (See Figure 4).

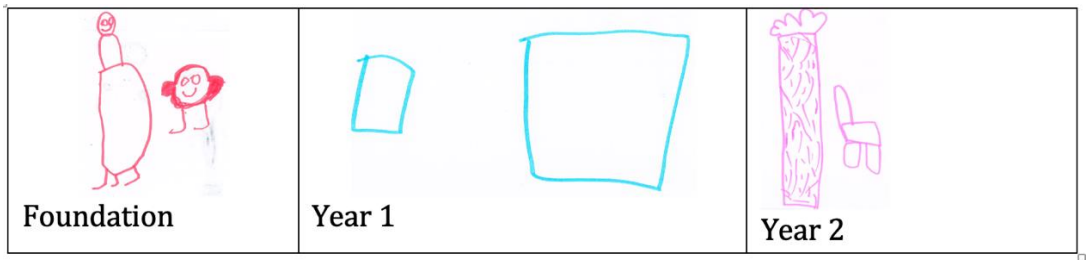


Figure 4. Examples of Category D: Figures not on a baseline

Category E: Figures on a baseline

32 drawings (29.4%) represented two (or multiple figures) with at least one distinguishably tall and another short. The figures tended to be recognisable objects that would logically be regarded as being tall or short, were drawn roughly in proportion to each other (See Figure 5).

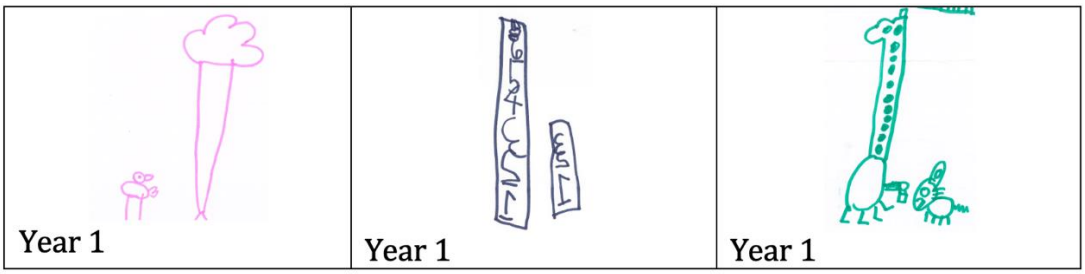


Figure 5. Examples of Category E: Figures on a baseline

A key feature of these drawings was the use of a common base line, or at least the apparent intention of alignment. The figures were often drawn at the bottom of the page, using the edge of the paper as the baseline. Clear representation was made of comparison and the difference in height rather than length in general; therefore, figures were drawn in vertical orientation. One student drew two columns of numbers (from 1 to 7 or 8, and the other 1 to 3) – perhaps depicting rulers for measuring length (See Figure 5, middle drawing).

Distribution of drawing types

The more sophisticated depictions of tall and short were more prevalent in Year 1 and 2 than in Preschool and Foundation. 67% of the children created a drawing of recognisable objects and made an obvious distinction between their heights. Of note is the persistence of incoherent drawings into Year 1.

Table 2: Number of drawings in each category for each year level

Drawing Category	Pre-school	Foundation	Year 1	Year 2	Total for category	Percentage n=109
A: Incoherent	7	8	6	0	21	19.3%
B: One Object	1	4	2	1	8	7.3%
C: Similar Size	1	3	1	2	7	6.4%
D: No Baseline	1	12	18	10	41	37.6%
E: Baseline	0	5	9	18	32	29.4%

DISCUSSION

In the five categories, we can see the parallel development of drawing skills and the communication of mathematical concept of height comparison. As expected from previous research (Carruthers & Worthington, 2005; Machón, 2013), most of the pre-school children were still exploring drawing as mark making (Category A). Of some concern, are the six Year 1 children (6-7 years) whose drawings were still in Category A: Incoherent, at a time when curriculum expectations involve the creation and interpretation of basic mathematical drawings and diagrams. Given the links that have been made between 5- to 6-year-old's ability to create schematised drawings and their later mathematical performance (Poland & van Oers, 2007), some appropriate teaching intervention to support drawing development seems warranted.

Also as anticipated, the majority of 6- to 8-year-olds had progressed to drawings as a representational medium for external objects which can be used to communicate mathematical concepts (Categories D & E). The task of drawing something tall and something short prompted the children to utilise schemas for proportion and for baselines. Interestingly, in the MacDonald (2010) study of 5- to 6 years, 89% of the children produced drawings of tall and short that clearly demonstrated "… comparing objects directly by placing one object against another" – which would coincide with Categories D and E in the present study. In contrast, in this study only about half of that age group (Foundation) produced comparison drawings. The MacDonald (2010) paper does not mention what the remaining 11% of 5/6-year-old children drew, making the interpretation of Category B: One Object, and Category C: Similar Height, somewhat isolated.

Across all year levels, Category B and C drawings were produced by 13.7% of the participants. Tall and Vinner's (1981) theory of concept images offers one interpretation. If a drawing of just one object (usually of something known to be tall in relation to a child) is accepted as the 'evoked concept image', then the drawings may indicate the presence of an emerging cognitive structure related to relative length, but

the second object of comparison has not been included. On the other hand, the drawings of two or more objects of similar height and size might be representations of 'evoked concept images' that show comparison, but not difference in length. A further interpretation of this situation may be that these children are in a transitional phase in the development of mathematical drawing and that teacher intervention with the purpose of highlighting features of effective representations would be very timely. In other words, teachers may be able to boost these children's drawing skills to better communicate their current mathematical understanding.

Limitations and further research

The participants in this study came from only one school, so further studies with a broader range of participants are needed to establish the applicability of the five tall/short drawing categories to different cohorts of children. Studies designed to probe the relationship between children's understanding of comparative length (or specifically height) and their ability to effectively communicate those understandings through their drawings could provide teachers with valuable advice about supporting children's development. This could be achieved by asking children to talk about their drawings to check for knowledge that has not been clearly represented in the drawing. Of particular interest would be intervention studies to test the proposition that children producing Category B and C drawings are demonstrating a readiness to progress to more effective mathematical drawings and that appropriate teacher instruction could prompt the transition.

CONCLUSION

The explication of the five categories of drawings directly addresses the research question: What types of drawings do 4- to 8-year-old children create to represent 'tall and short'? The findings tentatively suggest a developmental sequence in the representation of tall and short concepts through drawing, but the scattering of ages across the categories reminds us that children differ in the development rates. The simple task of asking children to draw something tall and something short can easily be replicated in other research but could also be used by teachers to focus their attention on the drawing development needs of their students, in addition to the conceptual learning needs for relative height.

REFERENCES

Arcavi, A. (2003). The role of visual representations in the learning of mathematics. *Educational Studies in Mathematics, 52*, 215-241.

Carruthers, E., & Worthington, M. (2005). Making sense of mathematical graphics: The development of understanding abstract symbolism. *European Early Childhood Education Research Journal, 13*(1), 57–79.

Bobis, J., & Way, J. (2018). Building connections between young children's representations and their conceptual development in mathematics. In Lai, M., Muir, T., & Kinnear, V. (Eds.). *Forging Connections in Early Mathematics Teaching and Learning* (pp. 55-72), New York; Springer. https://doi.org/10.1007/978-981-10-7153-9

Chigeza, P. & Sorin, R. (2016). Kindergarten demonstrating numeracy concepts through drawings and explanations: Intentional teaching with play-based *learning. Australian Journal of Teacher Education 41*(5), 65-77.

Ginsburg, H. P., Lee, J. S., & Boyd, J. S. (2008). Mathematics education for young children: What it is and how to promote it. *Social Policy Report of the Society for Research in Child Development, 22*, 3–23.

MacDonald, A. (2010). Young children's measurement knowledge: Understanding comparison at the commencement of schooling. In L. Sparrow, B. Kissane, & C. Hurst (Eds.), *Shaping the future of mathematics* education. (Proceedings of the 33rd annual conference of the Mathematics Education Research Group of Australasia). Fremantle: MERGA.

MacDonald, A., & Lowrie, T. (2011). Developing measurement concepts within context: Children's representations of length. *Mathematics Education Research Journal, 23* (1), 27-42.

Machón, A. (2013). *Children's drawings: The genesis and nature of graphic representation: A developmental study*. Madrid: Fibulas.

Papandreou, M. (2014). Communicating and thinking through drawing activity in early childhood. *Journal of Research in Childhood Education, 28*, 85-100.

Poland, M. M., & van Oers, B. (2007). Effects of schematizing on mathematical development. *European Early Childhood Education Research Journal, 15*(2), 269–293. DOI: 10.1080/13502930701321600

Tall, D., & Vinner, S. (1981). Concept image and concept definition in mathematics with particular reference to limits and continuity. *Educational Studies in Mathematics, 12*(2), 151-169.

van Oers, B. (1997). The narrative nature of young children's iconic representations: Some evidence and implications. *International Journal of Early Years Education, 5*(3), 273–245. DOI: 10.1080/0966976970050305

Way, J. (2018). Two birds flew away: The 'jumble' of drawing skills for representing subtraction pre-school to Year 1. In Hunter, J., Perger, P., & Darragh, L. (Eds.). *Making waves, opening spaces* (Proceedings of the 41st annual conference of the Mathematics Education Research Group of Australasia) pp. 98-101. Auckland: MERGA.

APPLYING A COMMOGNITIVE-BASED FRAMEWORK TO PROMOTE TEACHERS' COMMUNICATION ABOUT REASONING AND PROVING

Merav Weingarden and Orly Buchbinder

University of New Hampshire

This study presents how the commognitive-based Opportunities for Reasoning and Proving (ORP) Framework, developed for research purposes to analyze mathematical tasks, was applied as a learning tool for teachers. Seven novice secondary teachers, who participated in a professional learning community around integrating reasoning and proving, were introduced to the ORP Framework and engaged in a sorting tasks activity. We show how the ORP Framework helped teachers to focus on the ORP embedded in tasks, to attend to student mathematical work, and to communicate about ORP coherently and unambiguously. We discuss the affordances of using a framework, which relies on the operationalized discursive language of commognition, to promote teachers' communication around reasoning and proving.

REASONING AND PROVING IN MATHEMATICS TEACHING

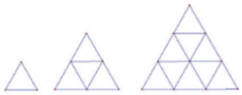
Mathematics educators and policymakers outline the vision of mathematics classrooms in which students develop proficiency with reasoning and proving (e.g., Hanna & de Villiers, 2012; NCTM, 2014). In this vision, teachers have a critical role in designing instructional activities that involve reasoning and proving, and teacher educators in preparing prospective teachers to design such activities (Buchbinder & McCrone, 2020; AMTE, 2017). However, what constitutes "reasoning" and "proving" has been an elusive topic (Reid & Knipping, 2010). For example, Stylianides (2008) defined "reasoning-and-proving" as a set of processes such as identifying patterns, making conjectures, and justifying, while others (e.g., Cirillo & May, 2020) focus on deductive reasoning and the logical structure of theorems and proofs. Jeannotte and Kieran (2017) argued: "what mathematical reasoning consists of is not always clear [and] it is generally assumed that everyone has a sense of what it is" (p. 1). Clarifying the notions of reasoning and proving in the school setting may aid teachers in providing their students with richer opportunities for reasoning and proving.

In our previous work (Weingarden et al., 2022), we utilized the discursive perspective of commognition to develop the *Opportunities for Reasoning and Proving* (ORP) Framework (described below) with which we conceptualized and operationalized the notion of reasoning and proving in mathematics classrooms by the opportunities provided to students to participate in certain types of discourses. In this study, we apply the ORP Framework in the context of teacher education and explore how it can support novice teachers in their communication about reasoning and proving.

2023. 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. 339-346). PME 46.

THEORETICAL PERSPECTIVE

Weingarden et al. (2022) developed the ORP Framework for characterizing mathematical tasks according to the ORP embedded in them. The framework draws on the robust theoretical tools of the commognitive perspective (Sfard, 2008), which views learning mathematics as a special type of discourse, and thus mathematical tasks can be characterized according to the type of discourse afforded to students by the task. The ORP in a task are determined by the objects at the core of the task: school-based (e.g., equation) or logic-based (e.g., conditional statement), and by the processes needed for solving the task: school-based (e.g., formulating an equation), logic-based (e.g., writing a conditional statement), or reasoning processes (e.g., generalizing, justifying). Table 1 presents the four types of ORP revealed in the previous study: Limited, Mixed, Logic-based, and Fully-Integrated ORP. The examples of tasks for each type of ORP are numbered in the order in which they were used in this study's intervention (see the Method section).

Tasks' characteristics	Examples [emphasis added]
	Limited ORP
Tasks that focus solely on school-based mathematical objects and include school-based processes.	3. *Solve* and *graph* the equation: $x^2 + 4x - 12 = 0$. 7. *Find a perimeter* of a rectangle whose one side is 5" and whose length is twice its width. 11. A farmer had some chickens and some cows. She counted 40 heads and 126 legs. How many chickens and how many cows were there?
	Mixed ORP
Tasks that involve the enactment of mathematical reasoning processes such as pattern identification, conjecturing, justifying, etc., on school-based mathematical objects.	1. *Create equations* that one can use to find the number of smaller triangles and the number of sticks for any given number triangle and *explain your reasoning*.  8. *Explain* how many solutions a quadratic equation can have. 9. *Make a conjecture* about the relationship between isosceles triangles and equilateral triangles and *justify* your thinking.
	Logic-based ORP
Tasks that are characterized by a logic-based object and can engage students with logic-based processes, such as identifying the hypothesis and conclusion of a given statement, or formulating the converse of a given statement.	2. *Underline the hypothesis* and *circle the conclusion* in the given statement: If both roots of quadratic function are positive then a>0. 4. Explain in your own words *what a counterexample is*. 6. Given a statement: A quadrilateral with two pairs of opposite congruent sides is a parallelogram. *Identify the hypothesis and the conclusion* of the statement and *determine if the statement is universal or existential*.

Tasks' characteristics	Examples [emphasis added]
	Fully-integrated ORP
Tasks that sensibly integrate both school-based and logic-based objects, such that students must operate on both, applying two types of processes: school-based and logic-based.	5. Come up with an example of a *conditional statement* that has to do with *linear functions* and equations and *determine whether the statement is true or false.* 10. *Prove or refute* the following statement: A quadrilateral with two pairs of opposite congruent sides is a parallelogram.

Table 1: The four types of ORP in mathematical tasks

With the ORP Framework providing a concrete operationalization for how reasoning and proving can be integrated in mathematical tasks, we hypothesized that it could be helpful for teachers aspiring to implement reasoning and proving in their classrooms. This assumption was anchored in two research strands. First, is a strand of research that explores the use of research-designed tools (e.g., observation protocols or teaching assessments) as pedagogical tools for teacher learning. For example, Candela and Boston (2022) examined how teachers using the Instructional Quality Assessment tool helped them to reflect on their practice and improve their teaching. The second strand relates to teachers' pedagogical discourse around learning and teaching. While pedagogical terms such as "high-level thinking" and "conceptual understanding" became ubiquitous in the discourse of teachers and teacher educators, their meaning and how it is manifested in mathematics classrooms often have been vague and elusive. Thus, the communication about these terms is often incoherent or ambiguous (Weingarden & Heyd-Metzuyanim, 2023). This ambiguity is also recognized with respect to reasoning and proving, as mentioned above. Thus, introducing teachers to the ORP Framework may be beneficial for creating a common language to talk about reasoning and proving. This paper begins to explore this assumption, and attends to the research question: *How does the ORP Framework contribute to novice teachers' communication around integrating reasoning and proving in their teaching?*

METHODS

This study is part of a larger project investigating how beginning teachers learn to integrate reasoning and proving in their teaching. The first stage of this project designed a capstone course *Mathematical Reasoning and Proving for Secondary Teachers* (Buchbinder & McCrone, 2020) and examined how prospective secondary teachers' (PSTs') expertise toward reasoning and proving develops as a result of their participation in the course (Buchbinder & McCrone, in press). The second stage of the project followed the PSTs, who took the capstone course, into a year-long supervised internship; and the third stage followed the same teachers for the first two years of autonomous teaching. During the third stage, the teachers participated in an online Professional Learning Community (PLC). The PLC met four times per year, each meeting lasting 90 minutes. One of these meetings was devoted to identifying ORP in

mathematical tasks. Eight teachers participated in this PLC meeting, which included three parts. First, the pre-ORP task sorting activity, where teachers worked in three groups on sorting 11 tasks (shown in Table 1) in any way they see fit and naming their categories. Sorting, modifying and characterizing tasks has been shown to be beneficial to teachers' professional learning (e.g., Swan, 2007). Second, we introduced the ORP Framework by describing and exemplifying the objects (school-based and logic-based) and the processes (school-based, logic-based, and reasoning) in a separate set of tasks, and introduced the four types of ORP (Limited, Mixed, Logic-based, and Fully-integrated). In the third part, the post-ORP task sorting activity, teachers sorted the same 11 tasks again, according to the types of ORP. Data includes the video-recording and the transcript of the PLC meeting, and the pre-ORP and post-ORP sortings made by each of the three groups. In the pre-ORP and post-ORP episodes, we identified what categories of tasks the teachers created, how they named the categories, what sorting criteria they used, and their dilemmas or disagreements. We analyzed and compared teachers' pre-ORP and post-ORP discourse, including how they talked about the tasks, what they focused on, and whether and how they referred to reasoning and proving.

RESULTS

Pre-ORP sorting task activity: Overlooking the logic-based ORP

In the pre-ORP sorting task activity, teachers mainly focused on the level of thinking required from students, the complexity of the tasks, the extent to which the tasks involved multiple solution paths or a factual answer, and other pedagogical elements such as whether the task belongs to beginning or end of a unit (Table 2 shows the sorting of each of the three groups). For example, group 1 (Diane and Olive) sorted the tasks according to the assumed level of thinking. The three task categories they created were: *low-level thinking, moderate-level thinking, and higher-order thinking*. The *low-level thinking* category included tasks 2, 3, and 7, which they described as "straightforward tasks," "do this tasks," and "plug-and-chug tasks." In contrast, the *higher-order thinking* tasks (# 1, 5, and 9), were assumed to require "independent thought," and "explorations," where students "actually need to think about it" rather than being "fed the answer." Olive and Diane's discourse and sorting categories did not attend to the ORP embedded in the tasks. For example, task 3, which asks students to solve an equation, and task 2, which asks to identify the hypothesis and conclusion in a conditional statement, were similarly classified as "straightforward tasks" in the *low-thinking* category. This type of sorting did not distinguish between the tasks' topic and nature, the ORP embedded in them, and student mathematical work around them. Specifically, this categorization completely overlooked the logic-based ORP of task 2. Olive's comment "that's a nothing question," suggests that she did not attend to the logic-based characteristics of the task and its importance in explicating the logical structure of arguments and proofs. When classifying task 6, Diane and Olive contemplated whether it belongs to "low-thinking" or not, since "it starts as like a simple task", similar to task 2, but on the other hand, "it takes it a little bit further than just identifying the hypothesis and conclusion." Eventually, they classified this task as

moderate-level thinking, since determining if a statement is existential or universal "requires more discussion than graphing an equation," (c.f., task 3) but did not specify what this "more discussion" involves, how these tasks are different in the mathematical work students need to do, and how these characteristics relate to reasoning and proving.

Group	Categories names and task numbers		
1: Olive, Diane	Low-level thinking (2, 3 ,7)	Moderate-level thinking (4, 6, 8, 10, 11)	Higher Order Thinking (1 ,5 ,9)
2: Bella, Nancy, Francesca	Direct Approach (3, 4, 8)	Exploration (1, 7, 9, 11)	Conditional Statements/Proving (2, 5, 6, 10)
3: Riley, Wendy	Direct (2, 3, 7, 11)	Explanation (4, 8 ,10)	Multi-step (1, 6, 9) Open-ended (5)

Table 2: Pre-ORP categories and task numbers by group

Group 2, in contrast, recognized that some of the tasks are proof-oriented. Right from the start Nancy said: "one thing that's starting to jump out at me is that there's a couple [of tasks] that look like they're all about conditional statements, like number two, five…" This led the group to create a category *conditional statements/proving* which included tasks 2, 5, 6, and 10. All these tasks included the word "statement," and had students prove a statement, identify its hypothesis and conclusion, or produce a statement. Group 2's teachers also suggested a category of tasks that "don't have anything about proving." This category was further split into *exploration* (tasks 1, 7, 9, and 11) and *direct approach* (3, 4, and 8). The *exploration* problem category included "open-ended," and "experimental" questions, where "students have to play around with and figure out," and "need a more solid explanation to back it up." The *direct approach* category included questions "that have just one answer," do not imply "multiple ways to do it," and explicitly state "what students have to do." When discussing task 4, which asks students to explain in their own words what a counterexample is, the opinions split. Nancy and Bella wanted to categorize it as a *conditional statement/proving* task because "you're finding a counterexample for a conditional statement," but Francesca thought it fits better under the *direct approach* category. She explained: "I feel like counterexample is something that could be put at any level… but it's not asking you to necessarily find a counterexample. It's just asking what it is." This dilemma, similar to group's 1 uncertainty regarding task 6, shows that by classifying tasks into high-level (e.g., exploration problems, higher-order thinking) and low-level (direct, low-level thinking), the teachers overlooked the added value of tasks like 2, 4, and 6, that although are straightforward and require a factual answer, are important for students making arguments and proving (logical-based ORP).

Post-ORP sorting task activity: Focusing on reasoning and proving

In the post-ORP sorting activity, the teachers' discourse changed. First, instead of talking about the tasks' characteristics (e.g., straightforward, open-ended, high level, fits to advance students, can be part of a summative assessment), the teachers turned to talk about the mathematical work students need to do. For example, Olive suggested that task 10 has fully-integrated ORP and explained: "because you have to use the logic stuff and the school math content." She then concisely listed the logic-based processes and the school-based processes students need to do in the task. Regarding the same task, Bella said: "You need logic because you need to know what prove or refute means. But you also need to know what a quadrilateral with two pairs of congruent sides is. So that would be fully-integrated." Like Olive, Bella also clearly referred to the objects and the processes embedded in the task, which includes both the school-based components ("what a quadrilateral with… is"), and the logic-based components ("what prove or refute means").

The second change identified in the teachers' post-ORP discourse is that it became less ambiguous, and more objectified and concise compared to the pre-ORP discourse. For example, during the pre-ORP activity, Francesca described the tasks in the *conditional statement/proving* category as: "you're doing something but without it being an exploration. But you're also not proving it." With these vague terms, she tried to capture the essence of the logic-based ORP type of tasks, that can be straightforward ("without exploration") and not require proving, but still related to conditional statements ("you're doing something"). On the contrary, in the post-ORP activity, when sorting task 2, which includes logic-based ORP, Francesca explained that students are "just underlining [the hypothesis] and circling [the conclusion], but they still have to have that logic of it." That is, Francesca clearly and explicitly indicated what students need to do (underlying, circling) and use ("the logic of it") to solve the task.

Similar observations were revealed in all other groups, where vague terms and reliance on feelings about the difficulty level of the task, were replaced with the precise language of the ORP Framework and meaningful sorting of tasks according to types of reasoning and proving activity expected from the students engaged with the task.

DISCUSSION

We examined how the ORP Framework, developed for research purposes, can be used as a learning tool for teachers in a professional development setting. Our findings show that the ORP Framework helped teachers to attend to the ORP embedded in the tasks. In the pre-ORP sorting activity, the teachers' discourse was subjectified and was lacking a unified and coherent language to describe students' mathematical work. The teachers used vague and ambiguous terms (e.g., "needs discussion," "play around," "doing something") and focused on general pedagogical aspects, such as level of thinking or task complexity. The teachers also attempted to characterize the tasks by the keywords (e.g., "explain," "find," or "conditional statement") rather than focusing on the conceptual, mathematical work students need to do in the task. In contrast,

teachers' post-ORP discourse was more objectified, coherent, and focused on student mathematical work, and components of reasoning and proving. The operationalized characterization of ORP also helped teachers develop an objectified way of talking about tasks, including what the task is about, and what students need to do to solve it.

We find this outcome interesting, because our teachers, although novices, were well familiar with proof-related tasks, having developed and enacted many such tasks as PSTs in the capstone course (Buchbinder & McCrone, 2020). Yet when it came to identifying the potential of a task to engage students with reasoning and proving, the teachers lacked the common unambiguous language for describing this potential. The ORP Framework, by relying on the discursive language of commognition, provided teachers with such a common language.

The advantage of commognition (Sfard, 2008) is that it enables operationalized communication about mathematics teaching and learning. However, communicating about teaching through the commognitive lens, especially with teachers, is not a straightforward process. The ORP Framework, similar to other tools and mediators developed based on commognition (e.g., Weingarden & Heyd-Metzuyanim, 2023), can help teachers to communicate about teaching more coherently without their familiarity and expertise in the commognitive framework.

The ORP Framework, developed first as a research tool for characterizing ORP in mathematical tasks (Weingarden et al., 2022), was found in this study, to contribute to teachers' emergent development of a common language (shared with teacher educators as well) for communicating about opportunities for reasoning and proving embedded in tasks. By this, our study contributes to the growing research on using research-based tools for teacher education and professional development (e.g., Candela & Boston, 2022). Moreover, by providing teachers with a coherent and objectified language to communicate about reasoning and proving, we step forward to support teachers' practices of identifying, designing, modifying, and enacting tasks that afford students ample opportunities for reasoning and proving – a need raised by many researchers and teacher educators (e.g., Hanna & de Villiers, 2012).

ACKNOWLEDGMENTS

This research was supported by the National Science Foundation, Award No. 1941720. The opinions expressed herein are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

Association of Mathematics Teacher Educators. (2017). *Standards for Preparing Teachers of Mathematics.* Available online at amte.net/standards

Buchbinder & McCone (in press). Preparing Prospective Secondary Teachers to Teach Mathematical Reasoning and Proof: The Case of the Role of Examples in Proving. To appear in *ZDM-Mathematics Education*

Buchbinder, O., & McCrone, S. (2020). Preservice teachers learning to teach proof through classroom implementation: Successes and challenges. *Journal of Mathematical Behavior, 58*, 100779.

Candela, A. G., & Boston, M. D. (2022). Centering professional development around the Instructional Quality Assessment rubrics. *Mathematics Teacher Educator, 10*(3), 204–222.

Cirillo, M., & May, H. (2020). Decomposing proof in secondary classrooms: A promising intervention for school geometry. In *Pro. of the 14th Int. Cong. on Math. Educ.*

Hanna, G., & de Villiers, M. (2012). Proof and proving in mathematics education: The 19th ICMI study. In *New ICMI Study Series* (Vol. 15). Springer.

Jeannotte, D., & Kieran, C. (2017). A conceptual model of mathematical reasoning for school mathematics. *Educational Studies in Mathematics, 96*(1), 1–16.

National Council of Teachers of Mathematics. (2014). *Principles to Actions: Executive Summary.* NCTM.

Reid, D. A., & Knipping, C. (2010). *Proof in mathematics education: Research, learning and teaching.* Sense.

Sfard, A. (2008). *Thinking as communicating*. Cambridge University Press.

Stylianides, G. J. (2008). An analytic framework of reasoning-and-proving. *For the Learning of Mathematics, 2*8(1), 9–16.

Swan, M. (2007). The impact of task-based professional development on teachers' practices and beliefs: A design research study. *Journal of Mathematics Teacher Education, 10*, 217–237.

Weingarden, M., Buchbinder, O., & Liu, J. (2022). Opportunities for reasoning and proving in mathematical tasks: A discursive perspective. In Lischka, A. E., Dyer, E. B., Jones, R. S., Lovett, J., Strayer, J., & Drown, S. (Eds). *Proc. 44th Conf. of the North American Chapter Int. Group for the Psychology of Mathematics Education.* Middle Tennessee State University.

Weingarden, M., & Heyd-Metzuyanim, E. (2023). What can the Realization Tree Assessment tool reveal about explorative classroom discussions? *Journal for Research in Mathematics Education, 54*(2)

FROM UNIVERSITY TO SCHOOL: EXPLORING BEGINNING TEACHERS INTEGRATING REASONING AND PROVING

Merav Weingarden and Orly Buchbinder

University of New Hampshire

We follow a beginning mathematics teacher, Olive, from the university-based course Mathematical Reasoning and Proving for Secondary Teachers *through the supervised internship where Olive taught in her cooperating teacher's classroom. By drawing upon Activity Theory, we compare her teaching within the two teaching settings, and we examine the opportunities for reasoning and proving she provided to her students in each teaching setting. As a prospective teacher, Olive provided her students opportunities for reasoning and proving. During the internship, these opportunities initially diminished due to institutional and contextual constraints. However, Olive gradually carved out unique paths to engage students with reasoning and proving as her teaching independence increased.*

BACKGROUND AND OBJECTIVES

The important role of teachers in supporting student engagement with reasoning and proving has long been recognized (NCTM, 2014). While there have been attempts to address the calls for enhancing teacher preparation around reasoning and proving within university programs (e.g., Buchbinder & McCrone, 2020; Conner et al., 2014), little is known about long-term development of proof-related practices of beginning teachers and the factors affecting it (Stylianides et al., 2017). Research suggests that transitioning from teacher preparation programs to supervised teaching experiences (hereafter, internship) is fraught with challenges. Beginning teachers often encounter tensions when balancing their commitments to the university and their mentor teachers, while also developing their own teaching styles and identity as mathematics teachers (Bieda et al., 2015; Smagorinsky et al., 2004). Stylianides et al., (2013) identified three specific challenges related to integrating reasoning and proving into mathematics teaching: beginning teachers struggle to implement proof-related tasks in real classrooms; have low knowledge of students' mathematical conceptions; and face non-productive classroom norms in mentor teachers' classrooms that do not promote exploring mathematical ideas, including reasoning-and-proving. However, novice teachers have been shown to hold on to some of the conceptual tools and practices developed during their training, and those tend to resurface by the second year of teaching (Grossman et al., 2000). Still, with respect to the complex practice of engaging students with reasoning and proving, there is limited knowledge about how prospective secondary teachers (PSTs) recontextualize what they learned in their teacher education program and apply it in their own mathematics classrooms. This study aims to contribute to this line of research by examining beginning teachers' transition from

2023. 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. 2, pp. 347-354). PME 46.

university to their internship, regarding implementing reasoning and proving, and the role of sociocultural contexts in shaping this transition.

THEORETICAL PERSPECTIVES

Situative perspective of teacher learning (Borko et al., 2000) is rooted in Vygotsky's sociocultural perspective that individual learning is a social phenomenon, affected by physical, social, and cultural contexts in which it occurs. The situative perspective places a critical role on teachers' identities which evolve across multiple and varied contexts such as one's own experiences as a learner, teacher preparation program, the internship site, and the school workplace (Thompson et al., 2013). In teacher preparation programs, PSTs encounter teaching practices aligned with ambitious, student-centered, discussion-rich instruction. However, the school where they teach may adopt more traditional instruction, and often explicitly or implicitly encourage beginning teachers to adopt the school's culture and norms. Thus, beginning teachers may experience tensions between the teaching paradigms across varied contexts (Feiman-Nemser, 2003; Thompson et al., 2013). Specifically, this applies to integrating reasoning and proving in teaching mathematics (Stylianides et al., 2013). Thus, it is reasonable to expect a dynamic, non-linear progression in the beginning teachers' development regarding the implementation of reasoning and proving from university-based teacher education programs to their internship sites and beyond.

To capture these dynamics, we utilize Activity Theory, specifically, Engeström's (1987) *collective activity system* model, to better understand the complexity of human learning from interactions with others, mediated by cultural tools and situated in social contexts. Increasingly, Activity Theory has been used to describe teacher learning in varied contexts (e.g., Ellis et al., 2019; Potari, 2013). Figure 1 shows Engeström's model overlayed with the description of its application to our study.

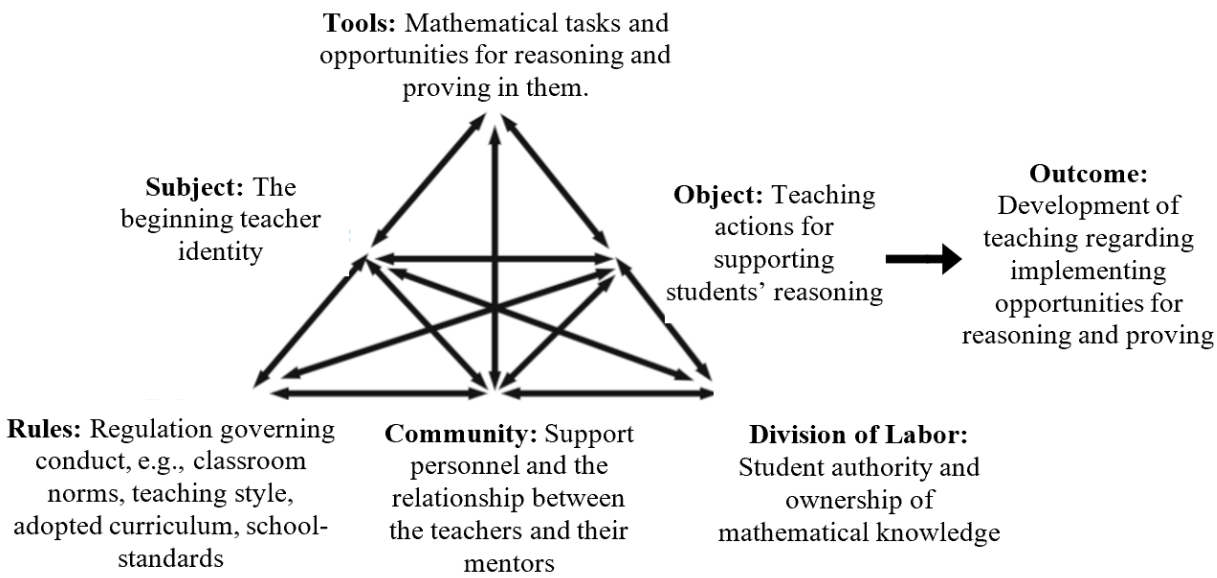


Figure 1: The collective activity system of activity theory applied to our study

This study is a part of the larger longitudinal project, which takes place in the US and examines how beginning teachers' knowledge, dispositions, and classroom practices related to reasoning and proving develop over time. Beginning teachers are followed for a one-semester capstone course *Mathematical Reasoning and Proving for Secondary Teachers* (Buchbinder & McCrone, 2020), a yearlong internship, and two years of autonomous teaching. In this paper, we focus on one beginning teacher – Olive, and compare her teaching, within an activity system, as a PST in the capstone course, and as an intern teaching collaboratively with her cooperating teacher. We examine how her teaching evolved in this transition. The overarching research question of this paper is: *How do Olive's activity systems as a PST and as an intern compare, and how does her teaching toward reasoning and proving develop across this shift?*

METHODS

Olive, a secondary mathematics education major, excelled in both mathematical and educational coursework. In Fall 2020 she completed the capstone course where she showed herself as an articulate and active participant, demonstrating positive dispositions towards proof and strong mathematical knowledge of proof. In this course, Olive designed and taught four proof-oriented lessons to a group of students in a local high school. After graduation, Olive started her supervised internship, gradually assuming teaching responsibilities in her cooperating teacher's high-school classroom. From each teaching setting, we collected 2-4 video-recorded lessons, including the lesson plans, written reflections (in the course), and post-lesson interviews.

Since Activity Theory was not developed specifically for mathematics education, we operationalized each node of the activity system (Figure 1) using additional analytic tools. The tasks (*Tools*) Olive enacted during her lessons were analyzed using the Opportunities for Reasoning and Proving Framework (Weingarden et al., 2022), which identified the types of opportunities for reasoning and proving embedded in the tasks. Olive's teaching actions (*Object*) were analyzed using three frameworks. The extent to which Olive's moves facilitate student reasoning (e.g., cueing, funneling) was analyzed by Teacher Moves for Supporting Student Reasoning (Ellis et al., 2019). The nature of Olive's questions (e.g., requesting a factual answer/ method/ idea) was analyzed by Teacher Support for Collective Argumentation framework (Conner et al., 2014). This analysis was overlayed with examining the *function* of Olive's question and moves based on Jeannotte and Kieran's (2017) Mathematical Reasoning for School Mathematics Framework, to capture the reasoning processes (e.g., validating, justifying) supported by these teaching actions. Student agency (*Division of Labor*) was identified by determining whether Olive's task enactment was student- or teacher-centered. Content analysis of Olive's reflections and interviews was used to glean findings about contextual factors, such as support personnel (*Community*), regulation governing conduct (*Rules*), and Olive's teacher identity (*Subject*), including her valued teaching actions regarding reasoning and proving, and the structural and institutional factors of her teaching settings.

RESULTS

Due to space limitations, we describe three lessons from the two teaching settings: one of Olive's lessons as a PST, and two of her lessons as an intern. Lesson one was in early October, the second lesson was about a month later, during Olive's "solo-week".

Olive as a PST

In Olive's lessons as a PST, Olive designed and enacted tasks that provide students with opportunities for reasoning and proving, such as identifying patterns, conjecturing, and proving (*Tools*). Figure 2 shows an example of an explorative task that Olive used in her first lesson as a PST.

A mathematician designed a figure using 3 **squares** sharing a vertex D. The brown square can be moved around a fixed point D. The mathematician noticed a relationship between some angles near the point D.

a) Can you spot the relationship she is referring to?

b) Make a conjecture about what happens when a pair of perpendicular lines is intersected by this right angle.

c) Let's prove the conjecture.

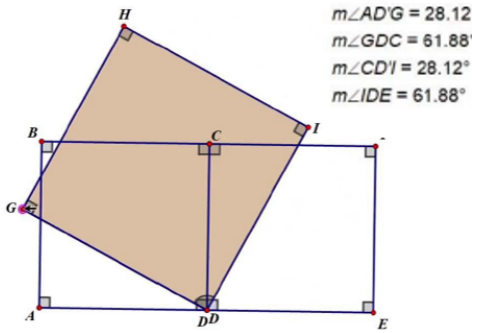


Figure 2: Olive's task in Lesson 1 as a PST.

This task asks students to explore a dynamic geometry sketch and to make a conjecture about the properties of two sets of perpendicular lines sharing the same intersection point and prove the conjecture. The task's set up, by giving students the time to explore and the opportunity to come up with a new narrative on their own, enables a high extent of student agency (*Division of Labor*). Olive's teaching actions in this lesson were mostly posing questions that requested ideas and inviting students to observe the relationship and generalize it. Other teaching actions included guiding, conceptual scaffolding, and directing student attention towards formulating the conjecture and then assisting them in completing the proof (*Object*). Overall, Olive's enacted teaching was characterized by rich opportunities for reasoning and proving provided to the students. In the capstone course, Olive was tasked with developing lessons focused on reasoning and proving (*Rules*) and was supported by her peers and encouraged by the instructor (the second author of this paper) to experiment, take risks, and show pedagogical creativity (*Community*). Reflecting on this lesson, Olive wrote:

> This theme was easy to integrate into a lesson plan. The content was properties of lines and angles. I was able to turn one of these properties into a situation for students to analyze, write a conjecture about, and prove. […] students had to eliminate excess information from a situation; they spent time devising a conjecture and determining what information was given and what needed to be proved. These are important features of proving.

This quote, as well as the other characteristics of the lesson, were typical of Olive's teaching during the course. Her discourse reflected confidence, comfort with designing proof-related activities, and recognition of their importance for student learning.

Olive as an intern

As an intern, Olive had to adhere to the norms of the school and follow the existing curriculum (*Rules*). Her cooperating teacher (CT), while generally supportive and welcoming, expected Olive to adhere to her teaching style and established classroom norms (*Community*). The class time was devoted to note-taking and computational practice. During the first weeks of her internship, Olive taught from the lesson plans developed by her CT and had limited input into the design and set up of the tasks (Lesson 1). With time, she embraced her growing independence and stepped into her solo-week teaching, where she was afforded to have more input into lesson planning and teaching, while still using her CT's overall unit plan and worksheets (Lesson 2).

Lesson 1. The lesson plan, designed by Olive's CT, was on solving equations. The tasks involved solving 12 linear equations, such as 11r+60=16 or 5(n+3)+9=3(n-2)+6. Olive looked for volunteers to solve the equations, but when no one did, she modelled solving the equations on the board, herself, while asking students about the next steps. Her questions included requesting facts ("What is positive 60 minus 60?") and requesting a method ("How am I gonna move the 11 away?"). Her teaching moves included mainly cueing ("Are 11 and r [11r] being added?", "If they're hugging like that, are they being added?"). Although students responded to Olive's questions, their agency was restricted, they were not given the opportunity to author new mathematical narratives (*Division of Labor*), and the task they engaged with was characterized by limited opportunities for reasoning and proving (*Tools*). Olive's teaching actions (*Object*) broke the task into procedural micro-tasks, like identifying operations or proving a factual answer (e.g., "What should I move first?", "my next step is combining like terms") and did not involve requesting an idea and pressing for reasoning.

In the post-lesson interview, Olive expressed frustration and dissatisfaction with the lesson designed by her CT (*Subject*). She described an envisioned alternative for this lesson where students explore the consequences of various operations on both sides of an equation. For example, whether adding 24 to both sides of an equation 4p=5 is a "helpful" operation, why or why not. This way students could explore more the reasoning and sense-making behind the rules for solving equations.

Lesson 2. This lesson was about inequalities and took place during Olive's solo-week. Olive planned the lesson based on her CT's notes and worksheets but added an "exploration," as she described it, to the lesson, where students had to discover what happens when an inequality is multiplied or divided by a negative number. Olive led the whole class discussion, where she invited students to write different inequalities and multiply or divide them by positive or negative numbers. She invited students to notice patterns and come up with conjectures. Students worked in pairs while Olive facilitated their thinking through guiding moves to help them find "a rule that always

works." For example, some students came up with a conjecture "when you divide or multiply by a negative number, the numbers will always be positive." In response, Olive offered a counterexample and guided the students toward validating their conjecture. Pointing to 72<104 divided by (-2) she said: "but when we divided it here, it became negative. Right? So, it's not always gonna become positive."

This task involved opportunities for reasoning and proving such as conjecturing, justifying, and validating (*Tools*). Olive's teaching actions (*Object*) included requesting ideas, directing by providing guidance, and engaging students in validating and conjecturing. The task set-up was student-centered resulting in greater student agency and authority by producing a new mathematical narrative (D*ivision of Labor*).

The rest of the lesson followed the CT's lesson plan, which like Lesson 1, contained strictly procedural tasks with no opportunities for reasoning and proving. When interviewed, Olive explained: "I wish there was more time to dedicate towards that (the exploration activity), but I needed to get through the lesson plan." Further, Olive explained that she introduced the "exploration" activity since she was worried about her CT's lesson plan presenting the rule of "flipping the inequality sign without any exploration." She said: "I just was very worried they (students) were going to ask why. I wanted them to see it for themselves and to understand why that was the case." Olive was also aware that this exploration does not constitute proof ("Of course, obviously an example doesn't constitute proof") but felt that it was more advantageous for students than merely receiving the rule from the teacher. Reflecting on this activity and her teaching more broadly, Olive said:

> They (students) did have opportunity to write, finish a conjecture and explore different examples. […] In my head they (students) are doing those explorations all the time... I think that that's how [students] learn and so, I would love to dedicate more time to specific things like that exploration. […] I always try to think of ways to incorporate it.

Analyzing Olive's discourse revealed that although she was expected to adhere to her CT's lesson plan (*Rules* and *Community*), she valued teaching actions and learning outcomes aligned with reasoning and proving practices (*Subject*), thus, she found a way to include proof-related activity in the lesson.

DISCUSSION AND SCIENTIFIC SIGNIFICANCE OF THE STUDY

Our analysis point to a trajectory of Olive's teaching development toward integrating reasoning and proving across two teaching settings. As a PST, the capstone course provided a rich teaching experience for Olive: she taught a full-length lesson to real students using her own lesson plan. Her enacted lessons (tasks, teacher moves and student agency) embedded rich opportunities for student engagement with reasoning and proving, suggesting emergent expertise for teaching mathematics via reasoning and proving (Buchbinder & McCrone, 2022). The relaxed rules and supportive community of peers contributed to Olive's developing teaching identity, confidence, and comfort with engaging students in reasoning and proving.

But although the course setting approximated classroom teaching, it did not model it entirely (Grossman et al., 2009): Olive did not have to follow a set curriculum nor respond to institutional considerations. As an intern, Olive found herself restricted by the school rules and culture, having to teach from her CT's lesson plans which provided students with limited, if at all, opportunities for reasoning and proving. This teaching style conflicted with Olive's developing teaching identity, as she repeatedly indicated.

It has been suggested in the literature that interns tend to abandon ambitious practices of their teacher preparation programs for the teaching practices of their CTs (e.g., Bieda et al., 2015). Our study provides evidence to the contrary: rather than adhering to her CT's procedural, drill-and-kill style, Olive held on to her desire to implement reasoning and proving. In her teaching, she actively sought ways to do so, while navigating the challenging institutional context (Herbst & Chazan, 2011).

In the case study methodology, it is critical to ponder: what is this a case of? Olive's self-selection to participate in this research project suggests a predisposition to embracing reasoning and proving as a teaching approach. Nevertheless, we treat this case as a proof of existence. We assert that Olive's case illustrates how beginning teachers' emergent teaching expertise coupled with productive beliefs about teaching mathematics via reasoning and proving may be retained, despite an unfavorable institutional context, and may begin to resurface in teaching practice over time. As we continue to follow Olive's first steps as an autonomous teacher, this study will provide additional insights into how this process unfolds.

ACKNOWLEDGMENTS

This research was supported by the National Science Foundation, Award No. 1941720. The opinions expressed herein are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

Bieda, K. N., Sela, H., & Chazan, D. (2015). "You are learning well my dear" Shifts in novice teachers' talk about teaching during their internship. *Journal of Teacher Education*, *66*(2), 150-169.

Borko, H., Peressini, D., Romagnano, L., Knuth, E., Willis-Yorker, C., Wooley, C., ... & Masarik, K. (2000). Teacher education does matter: A situative view of learning to teach secondary mathematics. *Educational Psychologist*, *35*(3), 193-206.

Buchbinder, O., & McCrone, S. (2022) Guiding principles for teaching mathematics via reasoning and proving. *Proceedings of the 12th Congress of the European Society for Research in Mathematics Education*. Bozen-Bolzano, Italy.

Buchbinder, O., & McCrone, S. (2020). Preservice teachers learning to teach proof through classroom implementation: Successes and challenges. *Journal of Mathematical Behavior*, *58*, 100779.

Conner, A., 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.

Ellis, A., Özgür, Z., & Reiten, L. (2019). Teacher moves for supporting student reasoning. *Mathematics Education Research Journal, 31*(2), 107-132.

Engeström, Y. (1987). *Learning by expanding: An activity-theoretical approach to developmental research.* Orienta-Konsultit.

Herbst, P., & Chazan, D. (2011). Research on practical rationality: Studying the justification of actions in mathematics teaching. *The Mathematics Enthusiast, 8*(3), 405–462.

Grossman, P. L., Valencia, S. W., Evans, K., Thompson, C., Martin, S., & Place, N. (2000). Transitions into teaching: Learning to teach writing in teacher education and beyond. *Journal of Literacy Research, 32*(4), 631-662.

Grossman, P., Compton, C., Igra, D., Ronfeldt, M., Shahan, E., & Williamson, P. (2009). Teaching practice: A cross-professional perspective. *Teachers College Record, 111*(9), 2065-2100.

Feiman-Nemser, S. (2003). What new teachers need to learn. *Educational Leadership, 60*(8), 25-29.

Jeannotte. D. & Kieran, C. (2017). A conceptual model of mathematical reasoning for school mathematics. *Educational Studies in Mathematics, 96*(1), 1-16.

National Council of Teachers of Mathematics (2014). *Principles to Actions: Ensuring Mathematical Success for All*. NCTM.

Potari, D. (2013). The relationship of theory and practice in Mathematics teacher professional development: An activity theory perspective. *ZDM-Mathematics Education, 45*(4), 507–519.

Smagorinsky, P., Cook, L. S., Moore, C., Jackson, A. Y., & Fry, P. G. (2004). Tensions in learning to teach: Accommodation and the development of a teaching identity. *Journal of Teacher Education, 55*(1), 8-24.

Stylianides, G. J., Stylianides, A. J., & Shilling-Traina, L. N. (2013). Prospective teachers' challenges in teaching reasoning-and-proving. *International Journal of Science and Mathematics Education, 11*(6), 1463–1490.

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.), *Compendium for Research in Mathematics Education* (pp. 237-266). NCTM.

Thompson, J., Windschitl, M., & Braaten, M. (2013). Developing a theory of ambitious early-career teacher practice. *American Educational Research Journal*, 50(3), 574–615.

Weingarden, M., Buchbinder, O., & Liu, J. (2022). Opportunities for reasoning and proving in mathematical tasks: A discursive perspective. In Lischka, A. E., Dyer, E. B., Jones, R. S., Lovett, J., Strayer, J., & Drown, S. (Eds). *Proceedings of the 44th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (PME-NA). Middle Tennessee State University. Nashville, Tennessee.

LEARNING ABOUT DIGITAL TECHNOLOGIES OF THE WORKING WORLD IN REGULAR MATH CLASSES? TEACHING COMPOSITE BODIES WITH 3D PRINT AS A LEARNING CONTEXT

Mira H. Wulff, Anika Radkowitsch and Aiso Heinze

IPN - Leibniz-Institute for Science and Mathematics Education, Kiel, Germany

A result of the digitalization of the working world is a change of competence profiles graduates need to enter the workplace. Consequently, relevant digital technologies should be integrated into the classroom providing opportunities for students to gain digital experiences. To facilitate teachers' acceptance of such instructional innovations, digital tools can be used as learning context instead of learning content. Following a design-based research approach, we developed a prototypical math unit using 3D print as learning context. Results of an evaluation with 101 German students in grade 5 to 7 indicate that the teaching unit has a significant effect on behavioural and cognitive components of students' 3D print-related self-concept and that the learning context of 3D print does not distract students from learning mathematics.

INTRODUCTION

The digital transformation of the workplace and consequently the rapid development of digital technologies leading to changes in professions is one of today's central challenges. Competence profiles of (future) employees are changing due to the transformation from mechanical or analogue to digital, automatized processes (e. g., Grundke et al., 2017). Consequently, (future) employees need different knowledge than before to be able to work with new machinery and technologies. In addition to subject-specific competences, this includes general digital-technological basic knowledge to ensure that students are able to enter the digitalized workplace. Therefore, not only extracurricular learning opportunities, but also innovative teaching concepts in general education schools are needed. An implementation of innovative concepts in regular general education is, however, only possible if both subject-related and workplace-related competences are focussed. Within the Danish-German DiASper project (Digital Working World from School Perspective) 3D print was chosen as an example of a relevant technology in the modern world of work that is connected both to the curricular contents of mathematics and the digital technologies of the workplace. The decisive factor for technologies collected under the umbrella term "3D print" is the differentiation from existing formative and subtractive production processes, resulting from the additive construction of an object (e. g., Barnatt, 2016). Goal of the project is to create learning opportunities for regular mathematics classes that facilitate an exploration of the 3D printing process, including an early vocational orientation, without interfering with the learning of mathematics. Using a design-based research

2023. 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. 355-378). PME 46.

(DBR) approach, a prototypical unit for mathematics class has been developed (Wulff et al., in press). This unit comprises seven geometry lessons and uses the modeling phase of the 3D printing process as a *learning context* for the *learning content* of composite bodies. The unit has been evaluated and progressively adapted after expert surveys with industrial representatives and mathematics teachers as well as in a laboratory context with 16 students.

This paper presents first results of the ongoing empirical evaluation study examining the effects of the prototypical teaching unit on students' attitudes towards 3D print as a digital tool in grades 5 through 7 in German secondary schools.

THEORETICAL BACKGROUND

3D print in mathematics education

According to the NCTM (2000), mathematics instruction in secondary education should, among other things, enable students to transition into the professional world. Using 3D printing in regular mathematics education can follow several approaches: The technology can be used as *learning content* (i. e., learning about the technology or materials needed for 3D printing), as a tool for *visualization* (i. e., visualizing (complex) mathematics for educational needs) or as *learning context* (i. e., focusing on subject-specific content in the context of 3D printing). Within the two main approaches in international research, 3D printing or parts thereof are mainly used as a didactical tool for visualization: Either using 3D printed objects for visualization aspects (e. g., Dilling, in press) or using the modeling aspect of the digital technology (i. e., the use of a CAD software) in geometry lessons (e. g., Lavicza et al., 2020) as well as art-related mathematical projects (e. g., Menano et al., 2019).

Students' attitudes: 3D printing related self-concept

According to Rosenberg & Hovland (1960), attitudes have three components: behaviour, affect, and cognition. Based on this division, Janneck et al. (2014), among others, defined a computer-related self-concept (i.e., attitudes related to one's own person in dealing with a computer), subdivided into a behavioural, an affective, and a cognitive component: The behavioural component describes the concrete experiences with the digital technology, the affective component encompasses emotional motives while dealing with computers and the cognitive component refers to the subjectively perceived competence and self-efficacy with regard to computers as well as strategies for dealing with the digital technology. Based on the computer-related self-concept defined by Jannek et al. (2014), the 3D printing related self-concept is for this paper understood as being composed of the previously mentioned three components. The affective and the cognitive components are both further divided into five subcomponents: positive feelings, fearfulness, understanding, designing, and tool perspective for the affective and 3D print related subjectively perceived competences, self-efficacy, internal attribution, external control beliefs, and strategies for the cognitive sub-components.

Within this paper, five of these sub-components are considered, as they are deemed of high relevance for a (future) engagement with said digital technology: *The behavioural component*, the affective sub-components (*positive feelings* and *fearfulness)* as well as the cognitive sub-components (*3D print related subjectively perceived competences* and *3D print related self-efficacy).* These sub-components combine all three components and it is assumed that, in addition to the possibility of direct confrontation with 3D printing and, thus, a build-up of action experience, an increase in positive feelings and a reduction in fearfulness of digital technology are of particular relevance for (future) engagement (Wakefield, 2015). Furthermore, the cognitive sub-components are focused, since perceived competence and self-efficacy affect behaviour (Bandura, 1978).

Curriculum contents related to composite bodies

The learning content *composite bodies* is part of the German mathematical curriculum for lower secondary education (MBWK, 2014). Students are expected to be able to deal mathematically with composite solids in a variety of ways. Thus, they should be able to construct composite bodies from basic solids and name basic solids in composite bodies. These constructions are to be executed both by drawing and digitally supported. In addition, computational aspects are to be focussed on: Equivalent to previous calculations on basic solids, students should be able to perform surface and volume calculations on simple composite bodies (i. e., didactically adapted composite bodies composed of a reasonable amount of previously known basic solids).

RESEARCH QUESTIONS

The present paper reports first results of the last step of the DBR process, that is, the evaluation of the developed teaching unit. The main goal of the unit is (i) to facilitate an examination of the 3D printing process including an early vocational orientation and (ii) to not neglect the learning of mathematical competences. For this purpose, an intervention study in a pre-post-design investigating the 3D related self-concept and competences regarding composite bodies was conducted.

Within the lessons, 3D printing itself was not the subject matter, but only a vehicle for mathematical learning. More specifically, students worked on various tasks to construct prescribed (composite) bodies by using the CAD software TinkerCAD. They were informed that (i) the goal of the lesson was to learn the curricular geometric content composite bodies and (ii) they would apply a professional CAD software for a digital construction process that corresponds to the modeling phase of the 3D printing process. Students constructed various composite bodies and also calculated the surface area and volume. Of course, in the end, students were allowed to start the 3D printer to print a 3D object they had designed according to their own ideas.

The following research questions were focused:

(**RQ1**) To what extent does the teaching unit on composite bodies using 3D printing as learning context influence students' 3D printing related self-concept?

(**RQ2**) To what extent does the teaching unit on composite bodies using 3D print as learning context affect students' mathematical learning of composite bodies?

Regarding RQ1, it was expected that students' fearfulness would decrease after the initial contact with the digital technology. Furthermore, an increase of the other components regarding students' self-concept was anticipated. With RQ2, we wanted to check whether the learning context of 3D printing interferes with the acquisition of competences of composite bodies. We expected that the learning context is not distracting from learning about the mathematical content within the unit.

METHOD

To answer these questions, $N = 101$ (58 f, 42 m, 1 n. n.) students in grade 5 to 7 ($n_5 = 23$; $n_6 = 57$; $n_7 = 21$) from four Northern German secondary schools answered two online surveys (data collection is ongoing). In between the two surveys, students were instructed in the unit on composite bodies using the learning context of 3D printing. In collaboration with each individual teacher, the unit was adapted to fit the learning groups' individual needs (e. g., more or less support for students).

Within the survey, students were questioned regarding the previously mentioned aspects of their 3D print related self-concept. To assess their 3D printing related self-concept, students were asked to evaluate 28 statements on a 4-point Likert scale ("disagree" to "agree"). Table 1 shows the internal consistencies of the five used subscales. A series of one-sided paired-samples *t*-tests were conducted to compare pre- and post-test data. Effect sizes were measured by Cohen's *d* for *t*-tests to show change in students' self-concept after the intervention.

Variable and example item	Item count	Reliability (pre-test)	Reliability (post-test)
Behavioural – action experience *(I have a lot of practical experience in working with 3D printers.)*	3	.62	.58
Affective – positive feelings *(3D printing technology fascinates me.)*	2	.75	.82
Affective – fearfulness *(I have inhibitions about using 3D printers.)*	3	.83	.86
Cognitive – 3D print related perceived competences *(I am more confident in using 3D printers than the average person my age.)*	3	.92	.92
Cognitive – 3D print related self-efficacy *(When I face problems in the 3D printing process, I find ways to solve them.)*	4	.86	.94

Table 1: Reliability coefficients Cronbach's Alpha (scales with three or more items) and Spearman Brown (scale with two items).

Furthermore, each survey contained three (pre-test) or four (post-test) respectively mathematical tasks related to the mathematical content to be fostered (naming geometric bodies of a composite body, surface area calculation, volume calculation). The mathematical tasks were used to check whether the mathematical content was learned during the teaching unit. For this purpose, one task per mathematical content and survey was created and used as an indicator in the pre- and post-surveys respectively. In task 1, a building was shown from four perspectives (front, right, back, and left). Students were asked to name the basic geometric bodies (e. g., prism, sphere, cylinder) of which the building was composed. In the pre-test four and in the post-test 5 bodies had to be named. To ensure comparability, the scoring was normalized (pre: 0.25 and post: 0.2 points per correctly named body). In task 2 and 3, respectively, students were given an inner-mathematical or contextualized task in which they had to provide the surface area or volume of a basic solid or a composite body. Table 2 gives an impression of the figures. For complexity reasons, students were asked to choose from eight given answers in an MC-design. Each of the eight answers could be achieved by a mathematical composition of the figure's given dimensions. Two values each corresponded and differed only in the dimension of the unit (e. g., 750 cm² vs. 750 cm³ for task 3 in the pre-test). Students got one point for the correct solution and 0.5 points if the number was correct but the dimension was wrong. Since the tasks are different in each survey, the responses are not directly comparable. A comparison with a control group is planned. Consequently, only a pre-post comparison within the experimental group is carried out for this paper.

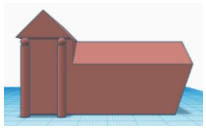
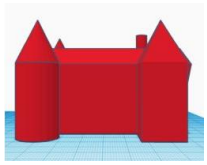
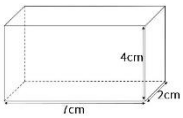
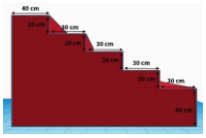
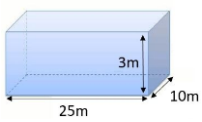
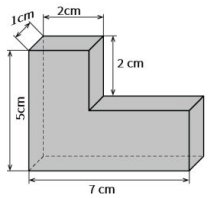
Task 1 (basic bodies)		Task 2 (surface area)		Task 3 (volume)	
Pre	Post	Pre	Post	Pre	Post
back view: 	right view: 				

Table 2: Sample images from the tasks used in the pre-test and post-test.

RESULTS

As shown in Table 3, results of paired-samples *t*-tests reveal a significant difference in the behavioural and cognitive components of the 3D print-related self-concept for all students between pre- and post-test. The differences correspond to small to medium effect sizes. The data also show a weak significant difference within students' positive feelings towards 3D print.

	Pre-test Mean (*SD*)	Post-test Mean (*SD*)	*t* (*p*)	Cohen's *d*
Behavioural – action experience	1.68 (0.60)	2.11 (0.70)	6.75*** (<.001)	.71
Affective – positive feelings	2.55 (0.93)	2.71 (0.80)	1.59# (.06)	.19
Affective – fearfulness	2.38 (0.89)	2.27 (0.86)	-1.16 (.13)	-.13
Cognitive – 3D print related perceived competences	1.87 (0.83)	2.21 (0.86)	3.28*** (<.001)	.39
Cognitive – 3D print related self-efficacy	2.06 (0.87)	2.35 (0.88)	2.66** (.005)	.34

Table 3: Comparisons of descriptive statistics and paired samples *t*-test of pre- and post-test scores (*** $p < .001$, ** $p < .01$, # $p < .10$).

As shown in Table 4, results of the paired-samples *t*-tests reveal a significant difference between pre- and post-test in all mathematical tasks with small effect sizes. However, there is a negative effect regarding the tasks related to volume calculation.

	Pre-test Mean (*SD*)	Post-test Mean (*SD*)	*t (p)*	Cohen's *d*
Naming geometric bodies	.61 (.26)	.73 (.19)	3.06** (.002)	.41
Surface calculation	.07 (.22)	.20 (.38)	2.64** (.005)	.31
Volume calculation	.53 (.43)	.37 (.41)	-2.70** (.004)	-.31

Table 4: Comparisons of descriptive statistics and paired samples *t*-test of Pre- and Post-test scores (** $p < 0.01$).

DISCUSSION

The present study reports first results of an evaluation study as a last step of a DBR project that developed a prototypical teaching unit on composite bodies connected to the learning context 3D print. The motivation of the project is the need for a basic preparation of students for digital workplace technology (Grundke et al., 2017; NCTM, 2000). The results in Table 3 show that the teaching unit positively affected the surveyed components of students' 3D printing-related self-concept (RQ1). This is relevant for later contact and engagement with the technology in the context of the working world (Wakefield, 2015). In summary, it can be said that students gain initial experience in dealing with 3D printing technology with the help of the teaching unit which positively affects their self-concept. It must be noted that actual skills in dealing with the 3D printer were not tested, but only the subjective individual attitudes of the

students to 3D printing. However, these are considered vital for future engagement (Bandura, 1978).

Since 3D printing was implemented as a learning context and not as learning content, the acquisition of competences related to the mathematical content of the unit is also of central importance and must not be neglected (RQ2). Nevertheless, our results suggest that students gained mathematical competences in two of three content domains (Table 4). This is although all three tasks within the post-test can be considered more difficult than the pre-test tasks from a mathematics-didactic perspective. In task 1, one more body had to be named for the normalized full score (1) than in the pre-test. In tasks 2 and 3 of the post-test, figures composed of cuboids (staircase and L-shaped body) were considered instead of cuboids (pre-test). This could explain the difficulty students had in completing the volume task in the post-test.

Limitations

Data for the present study were collected in the school year 2021/22, when schools were still reeling from the aftermath of school closures due to the Covid-19 pandemic. This may have had implications for students' prior knowledge regarding geometric solids. Furthermore, the unit was taught by the regular teachers. Although the unit was previously discussed with the teachers there may have been differences in the frequency of exposure to the technology and the 3D printer between the classes. Finally, in this preliminary analysis, the check of whether the learning context 3D print interferes with the learning of geometry relies only on pre-post-data of the intervention group and the comparison within pairs of different pre-post-tasks was used as a preliminary approach. The data collection is ongoing and data from a control group that received conventional instruction on the content of composite figures without 3D printing will be collected. Then a comparison of the pre-post-item pairs in an experimental design will be possible which yields more reliable results.

Outlook

Despite the limitations, the first promising results suggest a positive evaluation of the developed teaching unit in which students were able to gain experiences with 3D printing and to develop positive attitudes toward it, but which also does not neglect the mathematical content. We expect that these findings will be confirmed when analysing the full sample which also includes control group data. Within the project, further teaching units on mathematical content (e. g., plane representations, perspectives) were constructed within the design-based research project. In future steps, these will also be evaluated like the unit presented in this paper. In summary, we think that an implicit encounter with 3D printing as a learning context coupled with existing regular mathematical teaching content leads to a sustainable encounter with the digital technology without restricting mathematics learning. Using 3D print as a learning context (instead of a learning content) is an approach that (i) allows students to gain experiences with relevant workplace-related digital technologies and (ii) is more likely

accepted by mathematics teachers as it allows them to achieve the intended mathematical learning objectives in the classroom.

REFERENCES

Bandura. A. (1978). Reflections on Self-Efficacy. *Advances in Behavior Research and Therapy, 1*, 237-269.

Barnatt, C. (2016). *3D Printing: Third Edition*. California: CreateSpace.

Dilling, F. (in press). 3D-Printing in Calculus Education – Concrete Ideas for the Hands-on Learning of Derivatives and Integrals. In F. Dilling et al. (Eds.), *Learning Mathematics in the Context of 3D Printing*. Wiesbaden: Springer Spektrum.

Grundke, R., Jamet, S., Kalamova, M., Keslair, F., & Squicciarini, M. (2017). Skills and global value chains: A characterization. *OECD Science, Technology and Industry Working Papers 2017*(05), OECD Publishing.

Janneck, M., Vincent-Höper, S., & Ehrhardt, J. (2014). The computer-related self concept: A gender-sensitive study. *International Journal of Social and Organizational Dynamics in IT, 3*(3), 1–16.

Lavicza, Z., Haas, B., & Kreis, Y. (2020). Discovering Everyday Mathematical Situations Outside the Classroom with MathCityMap and GeoGebra 3D. In M. Ludwig et al. (Eds.), *Research on Outdoor STEM Education in the digital Age: Proceedings of the ROSETA Online Conference in June 2020* (pp. 23-30). Münster: WTM.

Menano, L., Fidalgo, P., Santos, I. M., & Thormann, J. (2019). Integration of 3D printing in Art Education: A Multidisciplinary Approach. *Computers in the Schools, 36*(3), 222-236.

Ministerium für Bildung, Wissenschaft und Kultur des Landes Schleswig Holstein (MBWK) (2014). *Fachanforderungen Mathematik. Allgemeinbildende Schulen, Sekundarstufe 1, Sekundarstufe 2.* Kiel: MBWK

National Council of Teachers of Mathematics (NCTM) (2000). *Principles and Standards for School Mathematics*. Reston, VA: National Council of Teachers of Mathematics.

Rosenberg, M.J. & Hovland, C.I. (1960) Cognitive, Affective and Behavioral Components of Attitudes. In: M.J. Rosenberg & C.I. Hovland (Eds.), *Attitude Organization and Change: An Analysis of Consistency among Attitude Components* (pp. 1-14). New Haven*:* Yale University Press.

Wakefield, R. L. (2015). The Acceptance and Use of Innovative Technology: Do Positive and Negative Feelings Matter? *The DATABASE for Advances in Information Systems 46*(4), 48-67.

Wulff, M. H., Wilken, M., & Heinze. A. (in press). 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 Classes. In: F. Dilling, et al. (Eds.), *Learning Mathematics in the Context of 3D Printing*. Wiesbaden: Springer Spektrum.

TEACHERS' MULTIPLE AND ADAPTIVE NOTICING DRIVEN BY THEIR FRAMING OF PROFESSIONAL OBLIGATIONS IN THE CONTEXT OF A PROVING ACTIVITY

Mei Yang, Andreas J. Stylianides and Mateja Jamnik

University of Cambridge

Research on teacher noticing can help us understand teachers' perception and cognition underlining their in-the-moment decisions, but this construct is seldom researched in the area of proof-related teaching. In this study we aimed to investigate patterns in teachers' ways of noticing in the context of a proving activity, and we tried to explain these patterns using teachers' framing of professional obligations. The data were drawn from semi-structured, vignette-based interviews with twelve Chinese pre-service and in-service secondary mathematics teachers. The findings showed that teachers used multiple ways to notice students' reasoning, and they adapted their ways of noticing to the validity of the presented student arguments. The possible mechanism of how their multiple and adaptive noticing was driven by their framing is discussed.

INTRODUCTION

Teacher noticing has been increasingly researched to explore teachers' perception and cognition underlining their decisions in teaching (van Es & Sherin, 2021). Yet, this construct is seldom studied in the area of proof teaching, which is central in mathematics education. Furthermore, research has often assessed teachers' expertise in noticing using a relatively large unit of analysis (e.g., assessing their overall performance in noticing a video clip), but this may obscure its dynamics. In science education research, Russ and Luna (2013) used a different way to study the dynamics of teacher noticing. They explored its local variation (i.e., a teacher's noticing varied across class discussion and lab time) and explained such variation by its relationship with teachers' dynamic epistemological framing (i.e., how a teacher framed what was learning in both situations). Inspired by their work, we aimed to explore the dynamic nature of teacher noticing in the context of a proving activity at a small grain size (i.e., looking at how teachers noticed each single student argument), trying to address two main questions: What patterns exist (if any) in teachers' ways of noticing and their framing of professional obligations in the context of a proving activity? To what extent can these patterns be explained by teachers' framing of their professional obligations?

THEORETICAL FRAMEWORK

Professional noticing

Research on teacher noticing (e.g., van Es & Sherin, 2021) generally focused on how teachers pay attention to and make sense of subjects ranging from student thinking to entire classroom situations. In this study, we focused on teachers' noticing of student thinking to explore the perceptual and cognitive processes underlining their responses

2023. 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. 363-378). PME 46.

to student thinking. Following Jacobs et al. (2010), we defined teacher noticing as consisting of three interrelated components: (1) selectively attending to noteworthy students' strategies; (2) interpreting students' understanding; and (3) deciding how to respond to students (differing from giving actual responses).

Framing of professional obligations

Teachers' framing, their sense of what is going on (e.g., what kinds of knowledge take place in a situation), influences their noticing (Russ & Luna, 2013). In this study, adapting Erickson and Herbst's (2018) notion of teachers' professional obligation, we conceptualized a specific type of teachers' framing – teachers' *framing of professional obligations* – to describe teachers' sense of what professional obligations they have in a situation. According to Erickson and Herbst's (2018) framework, we expected that teachers may potentially frame their obligations as obligations to the discipline of mathematics (e.g., to represent mathematical knowledge), the individual student (e.g., to serve students' needs to learn), the class of students (e.g., to support classroom interaction), and the institution of schooling (e.g., curriculum requirements, school schedule). Given the dynamic nature of framing, we also expected that teachers' framing of their professional obligations may be dynamic across different situations.

RESEARCH METHODS

Participants comprised twelve Chinese teachers, including four pre-service teachers, four novice teachers, and four experienced teachers, to reflect a diverse teacher profile. The novice and experienced teachers respectively had on average 2.75 years and 17.75 years of experience in teaching junior high school students aged 12-15. They were recruited through convenience sampling.

Using semi-structured individual interviews, we collected data from teachers online. Each interview lasted for one hour. Teachers were presented with a comic-style scenario which depicted a classroom proving activity. The design of the scenario was based on real-world classroom episodes described in prior studies and it reflected various elements of proof-related reasoning (see Yang et al., 2022 for more details). We designed the scenario in comic-style in order to make the information sufficiently realistic, but still abstract enough, aiming to draw teachers' attention to students' thinking in this proving activity (rather than other extraneous information such as the student's gender) and capture their interpretations in this context (Herbst et al., 2011).

The scenario consisted of eleven episodes. Each episode illustrated a student's argument during the proving activity. For example, in an episode, a student examined some examples and then confirmed that the conjecture was correct. Among the eleven episodes, five episodes depicted invalid arguments and six depicted valid arguments. After being presented with each episode, participants were asked to describe (i) what they paid attention to, (ii) how they interpreted the students' understandings, and (iii) in what ways they would respond to the students; these questions corresponded to the three components of teacher noticing. Then they were asked a follow-up question,

"Why would you respond to the students in this way?", from which we inferred what framing of professional obligations drove the teachers' decisions in that situation.

In this study, we considered each teacher's response to each student argument as the unit of analysis. This led to a total of 132 units. To analyse the data, we adapted a set of frameworks from previous research (e.g., van Es & Sherin, 2021; Lee, 2021; Furinghetti & Morselli, 2011), based on which we coded teachers' responses into different categories. Specifically, as to teachers' ways of attending, codes included: focusing on (1) the *outcomes* (e.g., conjectures), (2) the *mathematically superficial process* (e.g., the student examined more examples), and (3) the *essential process* (e.g., these examples were strategically identified with rationales) of students' reasoning. Note that the latter two ways were differentiated based on whether teachers mentioned mathematically important details of students' reasoning process. As to teachers' ways of interpreting, codes included: (1) the *descriptive* way to describe what teachers observed, (2) the *evaluative* way to evaluate whether students' reasoning was valid, and (3) the *interpretive* way to make inferences about students' reasoning (e.g., why students make this argument, etc). Regarding teachers' ways of deciding, we categorized them in three ways: (1) the *general pedagogical way* which was primarily related to pedagogy (e.g., organizing group discussion), (2) the *product-oriented way* that oriented students to produce products of proving (e.g., a proof), and (3) the *process-oriented responses* that oriented students to experience the process of proving (e.g., analysing why the conjecture was refuted). Based on the above-mentioned Erickson and Herbst (2018) framework, we coded teachers' framing of professional obligations as follows: *individual obligation*, *interpersonal obligation*, *disciplinary obligation*, and *institutional obligation*. Multiple coding was applied for responses that reflected more than one category. Based on the coding results, patterns of teachers' noticing and their framing were identified.

FINDINGS

Teachers' noticing

Teachers used on average 2.17 ways of attending, 1.97 ways of interpreting, and 1.39 ways of deciding per unit. This suggested that teachers were able to notice students' reasoning in a variety of ways, and they tended to use a mixture of ways to notice.

Table 1 shows the number and percentage of units which showed each way of noticing. Overall, teachers most often focused on the superficial process of students' reasoning (92%), interpreted their reasoning in an evaluative way (90%), and decided to respond in a product-oriented way (74%). Still, such high proportions did not exceed 100%, which means that sometimes teachers did not notice students' reasoning in the way they most frequently used. Meanwhile, all other ways of noticing were also identified (ranging from 17% to 69%). This means that sometimes teachers adopted ways they less often used to notice students' reasoning. Both pieces of evidence may suggest that teachers adapted their ways of noticing to the situation.

Table 1: Number of units (percentage of total units for type of argument) which showed teachers' different ways to notice students' arguments

Way of noticing		Invalid argument (60 units)	Valid argument (72 units)	Overall (132 units)
Attending:	Outcome	48 (80%)	43 (60%)	91 (69%)
	Superficial process	57 (95%)	64 (90%)	121 (92%)
	Essential process	28 (47%)	47 (65%)	75 (57%)
Interpreting:	Descriptive	38 (63%)	44 (61%)	82 (62%)
	Evaluative	56 (93%)	63 (88%)	119 (90%)
	Interpretive	21 (35%)	38 (53%)	59 (45%)
Deciding:	General pedagogy	9 (15%)	14 (19%)	23 (17%)
	Product-oriented	39 (65%)	59 (82%)	98 (74%)
	Process-oriented	39 (65%)	24 (33%)	63 (48%)

Multiplicity of teachers' noticing. To examine to what extent teachers used multiple ways to notice students' reasoning, we counted the number of ways (i.e., ranging from 0 to 3) they used in each unit, and calculated the percent of units in which they used different number of ways. As to the Attending aspect, teachers demonstrated a mixture of three ways in 37% of the units, and they used a mixture of two ways in 43% of the units. A similar pattern emerged in the Interpreting aspect. Teachers respectively showed a mixture of three ways and a mixture of two ways in 25% and 47% of the units. Although teachers used on average fewer ways of deciding than ways of attending and interpreting, they still used more than one way of deciding in 37% of the units. These results further indicated that when noticing a student argument, teachers were capable of using a variety of ways and they tended to use a mixture of ways to pay attention to, interpret, and decide how to respond to the student's reasoning.

Adaptability of teachers' noticing. The high amount of using an evaluative way (accounting for 90% of total units) suggested that teachers constantly evaluated the validity of student arguments. Then to what extent did teachers adapt their ways of noticing to the validity of student arguments? To address this question, we compared how often teachers used each way to notice valid and invalid student arguments.

The second and third columns of Table 1 listed the number and percentage of units which showed each way of noticing in both situations (i.e., the Invalid-argument situation and Valid-argument situation). Comparing data in both columns, three ways were more common in the Invalid-argument situation than in the Valid-argument situation: "Focus on the *outcome* of students' reasoning" (80% versus 60%), "Interpret in an *evaluative* way" (93% versus 88%), and "Decide to give a *process-oriented*

response" (65% versus 33%). Opposite patterns were observed for the three ways (i.e., "Focus on the *essential process* of students' reasoning", "Interpret in an *interpretive* way", and "Decide to give a *product-oriented* response"), which were more common in the Valid-argument situation. These patterns indicated that teachers may adapt the above six ways depending on whether the student argument is valid or not.

The proportions of the remaining three ways (i.e., "Focus on the *superficial process* of students' reasoning", "Interpret in a *descriptive* way", and "Decide to respond with *general pedagogy*") were similar in both situations. Besides, these three ways were more often used in combination with other ways rather than alone. This may suggest that teachers were less likely to adapt the use of these three ways to the validity of student arguments, and they tended to pair these three ways with other ways that were more dynamically adapted to notice students' reasoning.

Teachers' framing of their professional obligation

We identified a variety of teachers' framing of professional obligations from our data (see Table 2). In terms of their *obligations to individual students*, results showed that teachers mentioned four types of reasons that drove their decisions, including "to provide the student opportunities to explore how to prove" (8%), "to facilitate the student's thinking" (39%), "to point out limitations of the student's reasoning/methods" (17%), and "to solve the task" (40%). As to their *obligations to students' interpersonal interaction*, they mentioned relevant reasons in 6% of the total units. In terms of their *obligations to the discipline of mathematics*, teachers mentioned their obligations to "demonstrate rigour and the logic of mathematics" (26%) and "cultivate students' mathematics competence" (26%) in mathematics teaching. Teachers also mentioned their *obligations to the institutions* in 8% of the units, including their obligations to the department (e.g., curriculum requirement) and the school (e.g., class time), etc.

Multiplicity of teachers' framing of professional obligations. Similar to what we found in the patterns of teachers' noticing, results showed that teachers on average mentioned 1.70 types of professional obligations per unit. In half of the units, they mentioned more than one type of professional obligation. And in only 5% of the units did they explain their decisions as a personal preference (e.g., because of their teaching style) without mentioning any professional obligations. This suggested that teachers' framing of their professional obligations constantly played a role in driving their noticing. In parallel with the multiplicity of teachers' noticing, their framing of professional obligations in each unit was also multiple. In addition, our identification of teachers' framing was based on teachers' explanations of why they decided to respond to students in a certain way (i.e., an important component of teacher noticing). Therefore, the above results may suggest that teachers' hybrid ways of noticing were driven by their hybrid framing of professional obligations.

Adaptability of teachers' framing of professional obligations. Table 2 shows that a certain type of professional obligation was mentioned in some units, but not in other

units. Did teachers adapt their framing of professional obligations? To address this question, we compared how often teachers framed each type of professional obligation when student arguments were valid and invalid.

Table 2: Number of units (percentage of total units for type of argument) which showed teachers' framing of their professional obligations in different situations

Framing of professional obligation	Invalid argument (60 units)	Valid argument (72 units)	Overall (132 units)
Individual obligation (Explore proving)	6 (10%)	5 (7%)	11 (8%)
Individual obligation (Facilitate thinking)	30 (50%)	21 (29%)	51 (39%)
Individual obligation (Point out limitations)	19 (32%)	4 (6%)	23 (17%)
Individual obligation (Solve the task)	7 (12%)	46 (64%)	53 (40%)
Interpersonal obligation	1 (2%)	7 (10%)	8 (6%)
Disciplinary obligation (Rigour & the logic)	25 (42%)	9 (13%)	34 (26%)
Disciplinary obligation (Math competence)	15 (25%)	19 (26%)	34 (26%)
Institutional obligation	5 (8%)	5 (7%)	10 (8%)

Comparing the percentages of units in which teachers mentioned each type of professional obligations in both situations, we found that some types of framing were more common in one situation rather in another. When student arguments were invalid rather than valid, teachers more often mentioned their framing of individual obligations to *facilitate the student's thinking* (50% versus 29%) and *point out limitations of the student's reasoning/methods* (32% versus 6%), and their framing of disciplinary obligation to *demonstrate rigour and the logic of mathematics* (42% versus 13%). The pattern of the framing *to solve the task* stood out as different: the teachers mentioned it more often in Valid-argument situations than in Invalid (12% versus 64%). This suggests that teachers' framing is dynamic as is their noticing. They also adapted their framing of their professional obligations depending on the validity of student argument.

The patterns of the remaining four types of framing were similar between the two types of arguments. The disciplinary obligation to *cultivate students' mathematics competence*, which was mentioned relatively often regardless of the validity of student arguments, was more often mentioned in conjunction with other obligations rather than alone. This may suggest that while adapting other obligations to the validity of student arguments, teachers constantly kept this obligation in mind.

DISCUSSION AND CONCLUSION

To conclude, our results show that teachers used multiple ways to notice students' reasoning, and they adapted their ways of noticing to the validity of student arguments. These results provide us with a window into the possible mechanisms by which teachers' multiple and adaptive noticing is driven by their multiple and adaptive framing of professional obligations (see Figure 3).

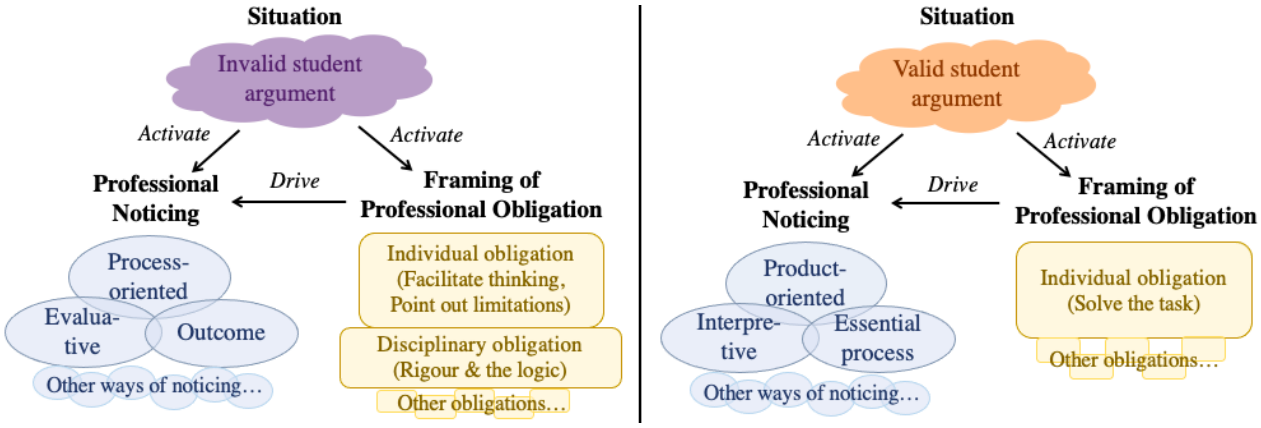


Figure 3: Mechanisms behind the multiplicity and adaptability of teacher noticing

Teachers more often framed their professional obligations in the Invalid-argument situations than the Valid ones so as to facilitate individual students' thinking, to point out limitations of the students' reasoning/methods, and to demonstrate the rigour and logic of mathematics. This may imply that when students raised invalid arguments, teachers tended to be concerned with how to guide students towards the "correct direction" according to disciplinary considerations. Under this framing of obligations, teachers tended to focus on the outcome of students' reasoning, assess which parts of their reasoning were incorrect, and decide what kind of process students should go through to get in the right direction.

By contrast, when student arguments were valid rather than invalid, teachers more often framed their professional obligations as to help students solve the proof task. This indicates that, in this situation, teachers tended to be concerned more with whether and/or how the task was solved by the students. Under such framing of obligations, teachers tended to focus on the essential process of students' reasoning (e.g., the strategies that students used), interpret students' reasoning in an interpretive way, and decide what possible products (e.g., a revised conjecture or a completed proof) students could produce in solving the task.

From Figure 3, we can identify differences between teachers' noticing and framing in both situations. Yet, this does not mean that teachers in one situation will not use the same ways of noticing or have the same types of framing as in another situation. Instead, results showed that teachers were capable of multiple ways of noticing and that such multiplicity can be explained by their framing of multiple professional obligations. Among these multiple ways of noticing or types of framing, some may be more common in one situation than in another, but they may still be activated across

situations. Also, some of them (e.g., the obligation to cultivate students' mathematical competence) may be constantly activated across both situations.

Given the dynamic natures of teacher noticing and framing, it is possible we may identify different patterns if the research context or participants change. Still, the mechanisms and patterns we identified in this study contribute empirically and theoretically to both the field of teacher noticing and proof-related instruction. First, the identified ways of noticing and framing of professional obligations enrich the field's understanding of teachers' perception and cognition in the context of proving activities, and offer researchers and teacher educators a starting point to further unpack the underlying mechanisms of teachers' decision making in this context. Second, they can be used as teacher training resources for teachers to disscuss and reflect on; introducing teachers to new ways of noticing or framing may transform their teaching. Third, future assessment and training of teacher noticing can consider its multiplicity and adaptability; for example, it can assess or aim to develop teachers' competence in using multiple and adaptive ways of noticing.

REFERENCES

Erickson, A., & Herbst, P. (2018). Will Teachers Create Opportunities for Discussion when Teaching Proof in a Geometry Classroom? *International Journal of Science and Mathematics Education*, *16*(1), 167–181.

Furinghetti, F., & Morselli, F. (2011). Beliefs and beyond: Hows and whys in the teaching of proof. *ZDM-Mathematics Education*, *43*(4), 587–599.

Herbst, P., Chazan, D., Chia-Ling Chen, Chieu, V.-M., & Michael Weiss. (2011). Using comics-based representations of teaching, and technology, to bring practice to teacher education courses. *ZDM-Mathematics Education*, *43*(1), 91–103.

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.

Lee, M. Y. (2021). Using a technology tool to help pre-service teachers notice students' reasoning and errors on a mathematics problem. *ZDM-Mathematics Education*, *53*(1), 135–149.

Russ, R. S., & Luna, M. J. (2013). Inferring teacher epistemological framing from local patterns in teacher noticing. *Journal of Research in Science Teaching*, *50*(3), 284–314.

van Es, E. A., & Sherin, M. G. (2021). Expanding on prior conceptualizations of teacher noticing. *ZDM-Mathematics Education*, *53*(1), 17–27.

Yang, M., Stylianides, A. J., & Jamnik, M. (2022, February). Chinese teachers' professional noticing of students' reasoning in the context of Lakatos-style proving activity. In *Twelfth Congress of the European Society for Research in Mathematics Education* (CERME12).

ELEMENTARY PRESERVICE TEACHERS' NOTICING OF EXEMPLARY LESSONS: A COMPARISON OF NOTICING FRAMEWORKS

Qiaoping Zhang[1] Yicheng Wei[1] Jing Liang[2]

1. The Education University of Hong Kong

2. Changchun Normal College

This study examined 111 mathematics preservice teachers (PSTs) noticing of exemplary mathematics lessons with different frameworks. The results showed differences in the content, stance and level of preservice teachers' noticing under the two frameworks. In the open framework, PSTs' noticing level is highly concentrated in Mixed(L2), with a wide range of noticing content on teachers' teaching strategies and skills. Teachers' noticing stance mainly uses explanations, most of which are detailed explanations of the noticing content. However, explanatory evidence is insufficient, and most strategies are not targeted. For the focused framework, PSTs' noticing levels are mainly located on Mixed(L2) and Focused(L3), with relatively concentrated content on students' mathematical thinking, knowledge and skills. Most of the noticing stances were descriptions. In addition, PG-group and UG-group have differences in noticing content, stance and level with the same framework. The findings can contribute to understanding the role of noticing frameworks in teacher noticing research and deepen our knowledge of Chinese PSTs' noticing ability.

INTRODUCTION

Teacher noticing refers to a teacher's ability to identify noteworthy instructional events from a classroom containing a large amount of information to interpret and respond appropriately (Bastian et al., 2022). This noticing ability is considered a core component of teachers' expertise (König et al., 2022). Ainley and Luntley (2007) described this expertise as attention-dependent knowledge, referring to skills that expert teachers use to attend to the cognitive and affective aspects of students' activities, and that guide the effectiveness of their instruction in response to what happens during a lesson. However, this is not an instinctive ability, which showed significant differences between preservice and expert teachers (Huang & Li, 2012; Stahnke & Blömeke, 2021). In previous studies, video-based methods were frequently used to examine PSTs and in-service experienced teachers' noticing skills (Amador et al., 2021). Using teaching videos in research can capture more details of the interactions between the teacher and students, and provide more opportunities for participating teachers to watch from multiple perspectives and reflect. However, with little experience in teaching and few interactions with students, PSTs always showed a low level of noticing and hardly perceived meaningful patterns from what they saw in the videos. They tend to have vague ideas about what they notice, and their explanations of the noticed content are often superficial. Some strategies have been

2023. 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. 371-378). PME 46.

proposed to develop PST' noticing skills, such as using reflective journals (van den Kieboom, 2021), teaching PSTs about clinical interview skills before noticing (Lee, 2021a), using technology-aided interventions (Lee, 2021b), Although these strategies did make a difference in enhancing PSTs' noticing ability. Still, they may not be suitable for PSTs to view a whole lesson. The difficulties in developing PSTs' noticing may be related to the theoretical framework in those studies. For example, Fisher et al. (2019)'s intervention experiment showed that the theoretical framework significantly facilitated the attending and explanation components. Numerous studies have always adopted the *Learning to Notice* framework to approach and guide teachers noticing (van Es & Sherin, 2010). Three prompts related to three skills of noticing (attend to, interpret, and decide to) were used to ask PSTs: "What did you attend to during the video?"; "How do you interpret the issues you attended to?"; and "How would you respond to the same situation?" . Since these questions in the noticing framework were general, the noticing level of PSTs' responses may still highly rely on their knowledge and teaching experiences. Some researchers suggested that PSTs could be explicitly provided with a set of focal points in the framework to guide their professional noticing (Lee & Choy, 2017). Inspired by Yang and Ricks (2012)'s three-point template for studying Chinese lesson study, Lee (2021b) used a three-point framework to enhance the noticing skills of US elementary PSTs. The *three-point* framework detailed how three interrelated focal points with which mathematics teachers design a lesson: the key point, the difficult point, and the critical point. The key point refers to the mathematical concept targeted in the lesson; the difficult point is the cognitive obstacle or stumbling block that students face when learning the key point; the critical point refers to the approach teachers can use to support students in their efforts to overcome the difficult point to learn the key point (Yang & Ricks, 2012).

The *Learning to Notice* and the *three-point* framework can provide valuable ways to examine what and how teachers notice when they observe lessons. The major difference between the two frameworks is that the former is more open and allows observers to select what they want to notice, while the latter is more structured and focused. *Learning to Notice* (the open framework) was developed in the Western classroom context, while the *three-point* framework (the focused framework) was developed in the Chinese context. In existing studies, the effects of using different theoretical frameworks on examining PSTs' noticing are relatively rare. Hence, the main purpose of this study was to explore the noticing abilities of Chinese elementary pre-service mathematics teachers and to compare the effects of the two frameworks. Specifically, we address the following two questions in this study: 1) What are Chinese PSTs' noticing of exemplary mathematics lessons; 2) What is the impact of the two noticing frameworks (open framework and focused framework) on PSTs' noticing?

THEORETICAL FRAMEWORK AND LITERATURE REVIEW

Mathematical teacher noticing

The concept of mathematics teachers' noticing consists of three practices—coding, highlighting and producing and articulating material representations (Goodwin, 1994).

Although teachers' noticing mainly involves two processes, namely attending to and making sense of particular events in an instructional setting, researchers have conceptualised noticing differently (Superfine et al., 2017). Based on their *Learning to Notice* framework, van Es & Sherin (2008) summarised that noticing skills for teaching mathematics consists of three main aspects: (a) identifying what is essential in a teaching situation; (b) using what one knows about the context to reason about a situation; and (c) making connections between specific events and broader principles of teaching and learning. Jacobs et al. (2010) conceptualised teacher noticing with three main components, including attending to children's strategies, interpreting children's mathematical understandings, and deciding how to respond to children's understandings. In this research, we consider noticing as a construct consisting of the three interrelated processes of attending to what is important or noteworthy in a classroom situation, interpreting and reasoning about the situation, and making instructional decisions based on interpretations.

Video-based approach to accessing teacher noticing

In the scoping review paper, most teacher noticing research instruments are considered as video materials from classroom practice (Weyers et al., 2023). Teaching videos have been used for decades in teacher learning, and it appears to show promise in supporting teachers in learning to notice (van Es & Sherin, 2006). By pausing and rewinding these videos, teachers can have more chances to focus on the events they care about, and to find supporting evidence for their claims. Also, videos provide a medium where teachers can critically analyse teaching practice in ways that are safely distanced from their own teaching experiences (Superfine et al., 2017). During the noticing engagements, the content that teachers notice and the complexity of their interpretations of what they notice influences their decision-making. Intervention therapy using instructional videos positively affects PSTs' professional noticing skills and content knowledge (Jacobs et al., 2010; Kosko et al., 2021). In this study, we used teaching videos and teachers' written responses as data sources to investigate mathematics PSTs' noticing. Stürmer and Seidel (2017) outlined three criteria for video selection: (a) teachers should perceive the videos as authentic examples of classroom practice, (b) the videos should serve to activate teachers' knowledge by being stimulating but not overwhelming, and (c) experts in the field should watch the videos as examples of the target practice. The mathematics teaching videos used in the project captured actual teaching by expert teachers in mainland China. These exemplary lesson videos were selected as viewing material to ensure the quality of the teaching being demonstrated. By watching these videos, PSTs could observe what expert teachers emphasised in their classroom teaching.

METHODOLOGY

Participants

A total of 111 PSTs was purposely recruited in one normal university located in the northeast of mainland China from April to May 2021. One group (UG-group, aged 20-21) consisted of 49 year-3 undergraduate students and the other group (PG-group, aged

22-23) included 62 year-1 master students. Participants in both groups majored in elementary education. They have learned mathematics courses and pedagogy courses. They were trained to be elementary mathematics teachers in the future. The difference was that the UG-group has completed teaching observations and didn't formally teach mathematics at schools, and PG-group hold bachelor's degree and completed teaching practicum during undergraduate study.

Theoretical frameworks used in the data collection

The open framework is very flexible which allows PSTs to select what they want to notice from the videos. The focused framework adapted from three-point framework (Lee & Choy, 2017) has a clear focus on what PSTs shall notice from the videos which may make more productive noticing. Besides considering three focal points, teachers also need to know students' pre-knowledge and cognitive thinking patterns, and further foresee their learning development (Tobias, 1994). We name it *Starting point*, for they are regarded as the very first step of lesson design (see Table 1). After random selection, 55 participants (24 in UG-group, 31 in PG-group) used the open framework and 56 participants (25 in UG-group, 31 in PG-group) used the focused framework.

	Attending to	Interpreting	Deciding to respond
Starting point			
Key point			
Difficult point			
Critical point			

Table 1. Focused noticing framework

PROCEDURE

Participants in these two groups were required to watch at one online exemplary mathematics lessons by using a noticing framework. These online lessons were

were published online by the Chinese Education Association (CEA) to provide mathematics teachers with professional development. All the lessons were taught by experienced and recognised elementary teachers. The contents of lessons included three main learning areas: Number & Algebra, Shape & Geometry, and Statistics & Probability. Each participant was required to watch at least one lesson and to select one of the noticing frameworks-the open framework or the focused Framework-to guide their observations. The total time for watching one complete lesson was around 40 minutes. After that, PSTs need to write down their observations.

DATA ANALYSIS

We analysed the valid written data using thematical analysis (Braun & Clarke, 2006). We first concentrated on two key aspects of noticing: what teachers notice and how they notice. Drawing from van Es and Sherin (2006)'s coding framework, what teachers attend to was examined through two dimensions, the Agent and the Topic that PSTs identified in the videos. How PSTs interpreted the events was analysed through the dimension of Stance. Stance consists of three categories: Describe, Evaluate, and Interpret. In addition to analysing what and how that PSTs noticed, the developmental

trajectory framework was adopted to evaluate their noticing levels (van Es & Sherin, 2010). There are four distinct levels of noticing from the lowest to the highest, these are Baseline (L1), Mixed(L2), Focused (L3) and Extended (L4). The baseline level indicates that teachers attend to events that cannot influence students learning directly. At the mixed level, participants primarily attend to teacher pedagogy or begin to attend to particular students' mathematical thinking; the focused level refers to attending to particular students' mathematical thinking; and the extended level refers to attending to the relationship between particular student thinking and teaching strategies.

RESULTS

What PSTs noticed of the lessons

As for the Agent of noticing, both UG-group and PG-group mainly focus on teachers and students. PSTs who used the open framework focused more on the agent of teachers and only a few (29.40%) on students, while PSTs who used the focused framework concentrated mainly on the agent of students. The main topics noticed by PSTs who focused on students were students' mathematical thinking, students' knowledge and skills, learning methods, mathematical myths, learning engagement, and interest in learning. As for the PSTs who focused on teachers, the main topics included teachers' mathematical thinking, teaching strategies, teaching skills, teaching beliefs, classroom management, teaching evaluation, teaching flow and teachers' gestures. Comparing the two frameworks, we can find that PSTs who used the open framework span a broader range of topics, and those who used the focused framework were relatively more concerned. This may be mainly because the four points in the focused framework help to guide PSTs to focus their thinking on some specific events. Specifically, in the starting point dimension, PSTs mainly focused on knowledge skills and teaching strategies; in the key point dimension, they mainly focus on knowledge skills, learning methods, and students' mathematical myths; in the difficult point, PSTs mainly focuses on mathematical myths, students' mathematical thinking, knowledge and skills, and learning methods; in the critical point, they mainly focused on teaching strategies, learning methods, and students' mathematical thinking.

How PSTs noticed

The UG-group had more evaluations than the PG-group when interpreting the events, while the PG-group used more descriptions and explanations. With the open framework, PSTs had most of the explanations in how to notice (72.13%). While with the focused framework, PSTs used more descriptions (51.80%). By analysing the four dimensions of the focused framework, it was found that both UG-groups and PG-groups used descriptions more often in four dimensions, especially in the key point and difficult point dimensions.

What are the levels of PSTS' noticing

Combined the analysis of noticing content and stance, it was revealed that both UG-group and PG-group' noticing skills levels were mainly on Mixed(L2). They primarily focused on teachers' teaching strategies, were able to highlight some meaningful events or details in their observations, and could provide certain specific events as

evidence to support their evaluations or explanations. However, these evidences were insufficient. They usually affirmed teachers' teaching strategies unilaterally or proposed alternative strategies, but they could not provide reasons for the strategies. PSTs in both groups were weak on Extended level (L4), which showed that they failed to reach high levels of noticing, and did not pay enough attention to the relationship between the students' thinking and teaching strategies. Comparing the two frameworks, it was found that the noticing level of PSTs who used the open framework was more at L2 (71.20%), while the noticing level of PSTs who used the focused framework was relatively evenly distributed at L2 (53.23%) and L3 (40.32%). PSTs noticing levels in the four dimensions of the focused framework were prominent at mixed level, especially the starting point dimension. UG-groups and PG-groups also showed some differences under the same framework. This is particularly evident in the focused framework, where PG-group only showed three levels (L2, L3, and L4) at different dimensions and most focused on Focused (L3). In contrast, UG-group could cover all four levels and mainly concentrated at L2.

DISCUSSION

In the study, the analysis of 111 Chinese elementary PSTs' noticing revealed that the overall noticing level of PSTs was concentrated on Mixed level (L2) and weak in Extended level (L4). PG-group focused more on students' learning and the UG-group focused more on teachers' teaching. When interpreting what they observed, the PG-group outperformed than the UG-group. PSTs' noticing stance was mainly to "describe", with the PG-group using "explain" more frequently than the UG-group. The finding echoes with previous research on teacher noticing and further shows that it's still a challenge for improving PSTs noticing, in particular at undergraduate level. It was found that the levels of *What* PSTs noticed and *How* PSTs noticed were not always consistent. For example, some PSTs who showed relatively low levels of noticing contents, mainly paied attention to teaching-related contents without mentioning students' thinking, but they could provide different specific strategies from multiple perspectives in the process of explaining. It is noted that these two stages (What teachers notice and How teachers notice) are inextricably linked, with the latter always based on the analysis of the former. For example, more focus on student thinking at the stage of *What* PSTs notice, consequently, more evidence-based details and interpretation of student thinking may be provided at the stage of *How* PSTs notice section. Hence, if we (teacher educators) aims to enhance PSTs noticing, we may need to broaden their views of classroom events and deepen their understanding of what happen in a mathematics classroom. These can be introduced in the pedagogy courses by theoretical instruction design or video-based case study.

PSTs always have little teaching experience or interaction with students, they tend to interpret students' thinking or teaching videos in vague ways, even not interpreting the work at all and only describing its surface-level details (Didis et al., 2016). The results showed that there were differences in PSTs noticing with the different frameworks. For example, PSTs using the open framework presented relatively broad topics and focused mainly on topics related to teachers' teaching. The PSTs using the focused framework

presented topics that were relatively focused on student mathematical thinking and other contents related to student learning. In the process of interpreting the content, the focused framework provides a structure to describe the four dimensions of thinking before and during the lesson, which in turn provides objective evidence for further analysis of the causes. Hence, compared to the open framework, the focused framework provides four major dimensions as cues for PSTs to conduct lesson observation, which help them to observe lessons more effectively.

CONCLUSION

The study not only analyzed PSTs noticing of exemplary lessons but also compared the effects of using different noticing framework on PSTs noticing level. In particular, this study incorporates the three-points framework originally developed in Chinese lesson study into Learning to Notice framework to specifically highlight what teachers notice during observing exemplary lesson videos. Hence, the findings provide theoretical and practical implications to the research field. First, compared to existing numerous studies on teacher noticing in western contexts, current finding could provide a cross-cultural perspective and empirical evidence to understand PSTs' noticing in the Chinese context. Second, the designed focused framework may contribute to develop research tools in teacher training programmes which aiming at improving PSTs' noticing skills. We acknowledge the limitation of small sample size and insufficient comparations with western partners. In future studies, we may examine a larger population on the effect of focused framework and conduct further research on whether the framework can improve teacher noticing in teaching practice.

Additional information

The research is support by Project R4300 and 04786 at the Education University of Hong Kong.

REFERENCES

Ainley, J., & Luntley, M. (2007). The role of attention in expert classroom practice. *Journal of Mathematics Teacher Education, 10*(1), 3–22.

Amador, J. M., Bragelman, J., & Superfine, A. C. (2021). Prospective teachers' noticing: A literature review of methodological approaches to support and analyze noticing. *Teaching and Teacher Education, 99*, 103256. https://doi.org/10.1016/j.tate.2020.103256

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.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101.

Didis, M.G., Erbas, A.K., Cetinkaya, B. et al. (2016). Exploring prospective secondary mathematics teachers' interpretation of student thinking through analysing students' work in modelling. *Mathematics Education Research Journal, 28*, 349–378.

Fisher, M. H., Thomas, J., Jong, C., Schack, E. O., & Dueber, D. (2019). Comparing preservice teachers' professional noticing skills in elementary mathematics classrooms. *School Science and Mathematics*, *119*(3), 142–149.

Goodwin, C. (1994). Professional vision. *American Anthropologist*, *96*(3), 606–633.

Huang, R., & Li, Y. (2012). What matters most: A comparison of expert and novice teachers' noticing of mathematics classroom events. *School Science and Mathematics, 112*(7), 420 – 432.

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.

König, J., Santagata, R., Scheiner, T., Adleff, A.-K., Yang, X., & Kaiser, G. (2022). Teacher noticing: A systematic literature review of conceptualizations, research designs, and findings on learning to notice. *Educational Research Review, 36*, 100453.

Kosko, K. W., Ferdig, R. E., & Zolfaghari, M. (2021). Preservice teachers' professional noticing when viewing standard and 360 video. *Journal of Teacher Education, 72*(3), 284–297.

Lee, M. Y. (2021a). The potential relationship between clinical interview skills and mathematics teacher noticing: An exploratory study. *International Journal of Science and Mathematics Education*, *19*(4), 793–814.

Lee, M. Y. (2021b). Improving preservice teachers' noticing skills through technology-aided interventions in mathematics pedagogy courses. *Teaching and Teacher Education*, *101*, 103301. https://doi.org/10.1016/j.tate.2021.103301

Lee, M. Y., & Choy, B. H. (2017). Mathematical teacher noticing: The yey to learning from lesson Study. In E. O. Schack, M. H. Fisher, & J. A. Wilhelm (Eds.), *Teacher Noticing: Bridging and Broadening Perspectives, Contexts, and Frameworks* (pp. 121–140). Springer International Publishing. https://doi.org/10.1007/978-3-319-46753-5_8

Stahnke, R., & Blömeke, S. (2021). Novice and expert teachers' noticing of classroom management in whole-group and partner work activities: Evidence from teachers' gaze and identification of events. *Learning and Instruction, 74*, 101464. https://doi.org/10.1016/j.learninstruc.2021.101464

Stürmer, K., & Seidel, T. (2017). A standardized approach for measuring teachers' professional vision: The observer research tool. In E. O. Schack, M. H. Fisher, & J. A. Wilhelm (Eds.), *Teacher Noticing: Bridging and Broadening Perspectives, Contexts, and Frameworks* (pp. 359–380). Springer International Publishing.

Superfine, A. C., Fisher, A., Bragelman, J., & Amador, J. M. (2017). Shifting perspectives on preservice teachers' noticing of children's mathematical thinking. In E. O. Schack, M. H. Fisher, & J. A. Wilhelm (Eds.), *Teacher Noticing: Bridging and Broadening Perspectives, Contexts, and Frameworks* (pp. 409–426). Springer International Publishing.

Tobias, S. (1994). Interest, prior knowledge, and learning. *Review of Educational Research, 64*(1), 37-54.

van den Kieboom, L. A. (2021). Using reflective journals to characterize pre-service teacher professional noticing skills. *The Teacher Educator, 56*(4), 347-371.

van Es, E. A., & Sherin, M. G. (2006). How different video club designs support teachers in "learning to notice". *Journal of computing in teacher education, 22*(4), 125-135.

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.

van Es, E. A., & Sherin, M. G. (2010). The influence of video clubs on teachers' thinking and practice. *Journal of Mathematics Teacher Education, 13*(2), 155–176.

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. https://doi.org/10.1016/j.tate.2022.103970

Yang, Y., & Ricks, T. E. (2012). How crucial incidents analysis support Chinese lesson study. *International Journal for Lesson and Learning Studies*, 1(1), 41-48.

EXPLORING THE ROLE OF PEDAGOGY IN MATHEMATICAL CREATIVITY VIA MULTIPLE SOLUTION TASKS: A COMPARATIVE STUDY OF TWO SCHOOLS IN CHINA

Ying Zhang

University of Cambridge

This study aimed to explore the relationship between pedagogy and mathematical creativity by comparing the mathematical creativity of students who experienced more student-centred pedagogy (SCP) with that of students who experienced more teacher-centred pedagogy (TCP). In total, 163 Grade 9 students from two schools in China, each enacting primarily one of these forms of pedagogy, participated in the study. Multiple solution tasks (MST) were used to measure mathematical creativity in terms of fluency, flexibility, originality. The total mean scores for fluency, flexibility, and originality of the SCP school were all higher than the respective scores of the TCP school though the differences were not statistically significant. Implications for research and practice are discussed in light of these findings.

INTRODUCTION

Creativity plays a crucial role in the full cycle of advanced mathematical thinking. Giftedness in mathematics assessments does not necessarily imply mathematical creativity (Sriraman, 2005), which requires divergent thinking that might not be covered in a national mathematics assessment. An example of this would be Chinese students who attained high scores in many international mathematics assessments but are in need of improvement in mathematical creativity (OCED, 2014). Relatedly, Lu and Kaiser (2022) conducted an empirical study among 107 Chinese upper secondary students, and found low levels of fluency and originality among participants reflecting their difficulties with attempting diverse ways to solve tasks. Regarding ways to foster mathematical creativity, some researchers suggested that, in contrast with teacher-centred pedagogy (TCP), student-centred pedagogy (SCP) has such a potential (e.g., Silver, 1997). Yet the relationship between pedagogy and mathematical creativity has attracted little research attention thus far. This study takes a step towards addressing this need for research by focusing on the following research question: Are there differences in mathematical creativity between students who have experienced more SCP and students who have experienced more TCP?

THEORETICAL PERSPECTIVES

The perspectives on TCP, SCP, and mathematical creativity are described as follows.

In a learning environment with TCP, the teacher primarily communicates to students through lectures solely designed to impart knowledge. SCP, in contrast, provides a learning environment where students construct their understanding, and teachers act as facilitators to “guide on the side” and help students achieve goals. While TCP is based

2023. 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. 379-386). PME 46.

on the behaviorist theory in which external stimuli causes behaviour changes, SCP involves constructivist and democratic principles where much knowledge is socially constructed (Serin, 2018). In this study, students categorized as having experienced more SCP did not necessarily only experience SCP-based lessons but did learn in an environment with more SCP compared to the comparative group.

The mathematical creativity in this study refers to the relative creativity, or creativity in school mathematics, which differs from that of professional mathematicians in that relative creativity is evaluated with reference to students' previous experiences and to the performance of other students who have a similar educational history (Leikin, 2009). Therefore, in this study mathematical creativity is to generate novel/original mathematical ideas, which are new to the person or the performance of other students in the similar educational history, with respect to the mathematics they have learned by discerning acceptable mathematical problems and models. I measured creativity in terms of three cognitive outcomes: fluency, flexibility, and originality. Fluency refers to the total number of appropriate problems generated by a solver; flexibility refers to the total number of strategies generated; originality refers to the uniqueness of one's solutions compared to others' response across two schools (Leikin, 2009).

METHODOLOGY

Comparative case study

In this study, Dulangkou Secondary School and School Y were selected as cases to play the role of SCP and TCP, respectively. Four Grade 9 classrooms, two from each school, were then randomly selected. In total, 83 Dulangkou students and 80 School Y students participated in this study. Importantly, Grade 9 is the third year of Chinese secondary education, and participants from the same school have received the same school instruction since Grade 7.

The two schools were selected due to having comparable features but different pedagogical approaches. Regarding similarities, firstly, both schools are in rural towns under the same county of the same city, so the schools follow the same educational policies and have similar economic conditions, though Y town has slightly better economic development and a better geographic location. Secondly, both schools randomly divide students into classrooms rather than dividing them based on achievement. Thirdly, both are the only school in their respective towns, both of which require recruiting students only from within the district; thus, both schools have similar sources for students. These similarities provide some control over confounding variables and allow for a meaningful comparison. Regarding the differences, School Y is one of the best-performing schools among all 14 towns and employs TCP, while Dulangkou is the most popular school among all towns due to its reformed SCP. Dulangkou has been using a reformed pedagogy for more than two decades and exemplifies the result of Chinese compulsory education reform (Sun & Wang, 2011).

I used the RTOP observation protocol (Piburn et al., 2000), which was developed to evaluate the extent to which a classroom adopts reform-based pedagogy, to verify that

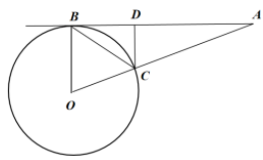
Dulangkou's mathematical pedagogy was more student-centred than that of School Y. The total RTOP score ranges between 0 and 100, in which lower scores reflect TCP environments and higher scores represent SCP environments. Specifically, the RTOP results showed that Dulangkou scored 21.4 higher on average than School Y. In total, four Dulangkou teachers (11 lessons) and five School Y teachers (13 lessons) were observed. All four participated classrooms were observed for at least three consecutive lessons, and the rest non-participated classrooms were randomly selected from each grade (Grade 7, 8, 9) for observation.

Due to the schools each having a distinct pedagogy, the class schedule also significantly differed between them. Specifically, Dulangkou students received eight mathematics lessons per week (45 minutes/lesson), while School Y students received seven mathematics teaching lessons and six mathematics self-study lessons (40 minutes/lesson) each week. Self-study lessons refer to the periods that students independently work on the assigned problems while the teacher simply monitors. Thus, each week School Y students received 160 more minutes of compulsory mathematics discipline than Dulangkou students.

Design of the multiple solution task

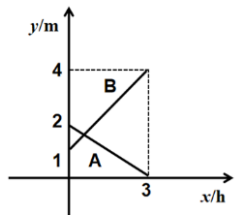
Solve the following problem in as many ways as possible.

Task A (Geometry)



The straight line *AB* is tangent to the circle with center *O* in point *B*. *OA* intersects the circle in point *C*. *D* is on *AB* so that *CD* is perpendicular to *AB (see figure)*. Prove that $\angle BCD = \angle BCO$.

Task B (Functional word problem)



A company has two cuboid reservoirs A and B. The water in reservoir A is injected into reservoir B at a speed of 6 m^3per hour. The functional relationship between water depth y (m) and injection time x (h) in reservoirs A and B is shown in the figure below. The functional relationship between the water depth y and the injection time x in the reservoirs A and B is: $y_A = -\frac{2x}{3} + 2$, $y_B = x+1$. After injecting $\frac{3}{5}$ hours, the depths of water in the two pools are the same. So, how long will the two reservoirs' water storage capacity be the same?

Figure 1. Tasks used in this study

This study used multiple solution tasks (MST) to indicate mathematical creativity. Two tasks, listed in Figure 1, were chosen for their coverage of geometry and functional word problems to assess multiple facets of students' conceptual and procedural knowledge. Task A is at an easy level of difficulty and Task B is at a difficult level. Circle related geometry and linear function are both important contents examined in

Zhongkao, the Chinese official Senior High School Entrance Examination held annually at the end of Grade 9.

Analysis method

Problems solved by students were first analysed based on appropriateness. The notion of appropriateness allows evaluating reasonable ways of solving a problem that potentially led to the correct solution outcome regardless of the minor mistakes made by a solver (Leikin, 2009). The data were then analysed based on fluency, flexibility, and originality. The detailed scoring scheme is explained in Table 1, which was adapted from Leikin (2009) in that students were given half credit for the partial correct procedure rather than an absolute zero or a full credit.

	Fluency	**Flexibility**	**Originality**	**Creativity**
Scores per solution	$Flu_i=1$ solution is appropriate $Flu_i=0.5$ solution is partial appropriate $Flu_i=0$ solution is inappropriate	$Flx_i=10$ solutions from a different group of strategies $Flx_i=1$ similar strategy but a different representation $Flx_i=0.1$ same strategy, same representation If $Flu_i=0.5$, $Flx_i=Flx_i/2$	$Or_i=10$ (P<15%) $Or_i=1$ 15%≤P<40 $Or_i=0.1$ (P≥40)	$Flx_i \times Or_i$
Total Score	$\sum_{i=1}^{n} Flu_i$	$\sum_{i=1}^{n} Flx_i$	$\sum_{i=1}^{n} Or_i$	$\sum_{i=1}^{n} Flx_i Or_i$

n is the total number of appropriate solutions, including partial appropriate solutions.

$P=(m_j/n)\cdot 100\%$, where m_j is the number of students who used strategy j.

Table 1. Creativity scoring scheme, adapted from Leikin (2009)

FINDINGS

Fluency, Flexibility and Originality

For Task A, School Y participants generated 81 appropriate solutions and 5 partial appropriate solutions; Dulangkou students generated 108 appropriate solutions and 1 partial appropriate solution. For Task B, Dulangkou produced 7 and School Y produced 8 appropriate solutions. Only six students from each school generated an appropriate solution(s) for task B, in which only one student from each school came up with two different strategies.

Table 2 and 3 demonstrate the number of students (n) generating the corresponding strategy, in which strategy A comprises of four different sub-strategies (A1, A2, A3, A4). For Task A, Dulangkou's solutions comprise of ten subcategories and School Y covers eight subcategories. For Task B, Dulangkou's solution comprise of two categories and School Y covers three categories. The value in the brackets represents the percentage of students within their schools' participants. For example, 14.4% Dulangkou participants and 6.3% School Y participants used strategy C. The corresponding originality is also described in both tables. In total, 21.7% Dulangkou and 15% School Y participants generated original solutions. Dulangkou have more participants who were able to generate original solutions among these two tasks.

Table 2. Distribution of categories of solutions solved by participants (Task A)

# of students / Strategy	Dulangkou (n = 83)				School Y (n = 80)				Originality
Strategy A	A1	A2	A3	A4	A1	A2	A3	A4	0.1 for 1.1 to 1.3; 10 for 1.4
	59 (71%)	5 (6%)	2 (2.4%)	2 (2.4%)	41 (51%)	0	3 (3.8%)	0	
Strategy B	26 (31.3%)				31 (38.8%)				1
Strategy C	12 (14.4%)				5 (6.3%)				10
Strategy D	1 (1.2%)				2 (2.5%)				10
Strategy E	1 (1.2%)				1 (1.3%)				10
Strategy F	1 (1.2%)				1 (1.3%)				10
Strategy G	1 (1.2%)				0				10
Strategy H	0				1 (1.3%)				10

Table 3. Distribution of categories of solutions solved by participants (Task B)

	Dulangkou	School Y	Originality
Strategy I	5 (6%)	2 (2.5%)	0.1
Strategy II	2 (2.4%)	4 (5)	0.1
Strategy III	0	1 (1.3%)	10

As indicated in Table 4, The total mean fluency, flexibility, and originality of Dulangkou participants were all higher than the respective scores of School Y students. However, the Mann-Whitney U test suggests such differences are not statistically significant, with the p-value of 0.105, 0.115, 0.301 for fluency, flexibility, and originality, which attributes to the small sample size, the scoring scheme's nature, and the difficulty of Task B.

Table 4. Mean and SD of Total Creativity for both tasks

	Fluency	Flexibility	Originality	Creativity
	mean±SD	mean±SD	mean±SD	mean±SD
Dulangkou	1.38±1.17	13.00±10.86	2.68±5.35	22.38±43.17
School Y	1.14±1.21	10.68±10.62	1.93±4.63	17.34±42.44

Creativity

All participants were ranked based on their total creativity score. 28% Dulangkou and 35% School Y participants obtained a zero score, and thirteen out of the top twenty performers are Dulangkou students. Specifically, for Task A, eight out of the top ten performers are Dulangkou students. However, according to Mann-Whitney U test, such differences are not statistically significant (p-value: 0.248).

Table 5. Creative Thinking Level, Adapted from Tatag (2011).

Level	**Characteristic of Creative Thinking Level**
Level 0 (Not Creative)	Students were not able to show any components of creativity ($Cr = 0$)
Level 1 (Almost Not Creative)	Students were able to show fluency without or with low originality and flexibility in solving problem ($Flu_i>0$, $Ori_i = 0.1$ and $Flx_{i\neq1}< 10$)
Level 2 (Quite Creative)	Students were able to show flexibility or originality in solving problem with low fluency ($0 < Flu_i\leq1$, $Ori_i > 0.1$ or $Flx_{i\neq1}= 10$)
Level 3 (Creative)	Students were fluent and then they were flexible or demonstrate originality ($Flu_i>1$, $Flx_{i\neq1}= 10$ or $Ori_i=10$)
Level 4 (Very Creative)	Students satisfied all components of creativity ($Flu_i>1$, $Flx_{i\neq1}=10$ and $Ori_i=10$)

To further analyse the two schools' performance, I grouped students' responses into Creative Thinking Level (CTL), which is the level I adapted from Tatag (2011) to classify students' creativity with relevance to the rest of the group. The indicators of CTL are explained in Table 5, and the CTL results are indicated in Table 6. The following ordinal logistic regression model was employed: logit $(P(Y \leq k|S)) = \log_e(\frac{P(Y\leq k|S)}{1-P(Y\leq k|S)})$, where Y denotes the CTL (k = 0, 1, 2, 3, 4) and S denotes the school (0 for School Y, 1 for Dulangkou). The model suggests that the odds of Dulangkou

students obtaining a higher CTL for Task A is 1.455 times, and for Task B is 0.959 times, as large as it is for School Y students. The estimated differences between the two schools are smaller for Task B than for Task A.

Table 6. Distribution of CTL

# of students	Level 0	Level 1	Level 2	Level 3	Level 4
			Task A		
Dulangkou	23 (27.7%)	26 (31.3%)	0	19 (22.9%)	15 (18.1%)
School Y	28 (35%)	21 (25%)	8 (10%)	16 (20%)	7 (8.8%)
			Task B		
Dulangkou	77 (92.7%)	5 (6%)	0	1 (1.2%)	0
School Y	74 (92.5)	5 (6.3%)	0	0	1 (1.3%)

Inter-rater Reliability

Two raters independently coded at least 12% of the student responses from each task of each sample. The inter-rater agreements were 96.4%, 91.1%, 96.4%, and the Cohen's Kappa was 0.819, 0.741, 0.819 for fluency, flexibility, and originality.

CONCLUSIONS AND DISCUSSION

The analysis of the MST performance showed that the SCP participants outperformed the TCP participants in terms of the total creativity; however, due to the small sample sizes and 93% zero-achievers in Task B, the differences are not statistically significant. School Y students receiving 160 minutes/week more compulsory mathematics learning time might narrow the differences between the schools' MST performance, which remains a possible avenue for future exploration. Overall, our study results implicated that integrating SCP into secondary mathematics might avoid excessively subjecting students to intense discipline while achieving a similar/better level of mathematical creativity. Moreover, 7% fewer zero achievers among Dulangkou participants indicates that SCP might help improve problem solving among the low performers.

Students' performance on Task B suggests two things. First, the benefit of SCP when solving easy-level MST might exceed the benefit from solving difficult MST. Second, when tasks are too difficult, students' mathematical creativity via MST can be restricted to their mathematical knowledge and problem-solving skills. As a result, the valid data received would be insufficient, and individual results could dominate the group results. I therefore suggest tasks aiming to indicate creativity be set at an easy or moderate level so that students' divergent thinking is not submerged. Although this study cannot guarantee the definitive causal claim between pedagogy and mathematical creativity, which attributes to the limitation of comparative studies: the quandary of "many variables, small-N" (Lijphart, 1971), it blazes a new path in this direction and underscores the need for more inquiry into this line of research. Future

studies should increase school cases and further control pedagogy as the main variable to robustly investigate the relationship between pedagogy and mathematical creativity.

ACKNOWLEDGEMENT

This work is based on Ying Zhang's PhD research at the University of Cambridge under the supervision of Professor Andreas Stylianides. A sincere gratitude to Guixi Liu and Ziang Zhang for their contribution as the second rater, and to Professor Ergul Demir and Ziang Zhang for their statistical knowledge and commentary.

REFERENCES

Leikin, R. (2009). Exploring mathematical creativity using multiple solution tasks. *Creativity in mathematics and the education of gifted students*, 9, 129-145.

Lijphart, A. (1971). Comparative politics and the comparative method. *American political science review*, 65(3), 682-693.

Lu, X., & Kaiser, G. (2022). Creativity in students' modelling competencies: conceptualisation and measurement. *Educational Studies in Mathematics*, 109(2), 287-311.

OECD. (2014). *PISA 2012 results: Creative problem solving: Students' skills in tackling real-life problems (Volume V)*. Pisa: OECD Publishing.

Piburn, M., Sawada, D., Turley, J., Falconer, K., Benford, R., Bloom, I., & Judson, E. (2000). Reformed teaching observation protocol (RTOP) reference manual. Tempe, Arizona: Arizona Collaborative for Excellence in the Preparation of Teachers.

Serin, H. (2018). A comparison of teacher-centered and student-centered approaches in educational settings. *International Journal of Social Sciences & Educational Studies, 5*(1), 164-167.

Silver, E. (1997). Fostering creativity through instruction rich in mathematical problem solving and problem posing. ZDM, 29(3), 75

Sriraman, B. (2005). Are giftedness and creativity synonyms in mathematics?. *Journal of Secondary Gifted Education*, 17(1), 20-36.

Sun, H., & Wang, X. (2011). From the Bottom Up: Effective School Improvement in Dulangkou Junior Secondary School in China. *Profesorado, Revista de Currículum y Formación del Profesorado*, 15(3), 51-65.

Tatag, Y. E. S. (2011). Level of students' creative thinking in classroom mathematics. *Educational Research and Reviews*, 6(7), 548-553.

SIXTH GRADERS' LEARNING OF MULTIPLICATIVE STRUCTURE PROBLEMS THROUGH THE VARIATION PRINCIPLE

Cristina Zorrilla[1], Ceneida Fernández[1], Anna-Katharina Roos[2], Salvador Llinares[1], and Susanne Prediger[2]

[1]University of Alicante, [2]TU Dortmund University

This study examines how sixth graders learn to solve multiplicative structure problems when they participate in an instruction focused on two connections: connecting multiplicative structure word problems and connecting the invariant structures from natural numbers to fractions. In a previous study, we identified five different students' modelling pathways. We extend this previous work by examining the changes in students' modelling pathways along the varied problem sets designed. The results show differences in the students' learning and shed light on the role of the variation principle in overcoming students' difficulties in solving this type of problems.

INTRODUCTION

Multiplicative structure word problems have been studied since the 1980s (overview in Greer, 1992); and students' difficulties regarding the problem structure and the type of numbers have been identified (e.g., Fischbein et al., 1985). The same situation (e.g., 15 cakes grouped in five boxes with three cakes each) can be transformed in three types of multiplicative structure problems according to its unknown (Greer, 1992): multiplication (the total amount is unknown, i.e., total of cakes), partitive division (the quantity per group is unknown, i.e., cakes per box), and quotitive division (the number of groups is unknown, i.e., number of boxes). However, many students deal with problem structures in isolated ways rather than making these connections of inverse relationships. Furthermore, previous studies have shown that also students who successfully solve these problems with natural numbers, do not succeed for rational numbers (Fischbein et al., 1985; Greer, 1992).

In our design research study, we investigated how to engage students in two types of connections to overcome these difficulties: (i) connecting multiplicative structure problems, i.e., multiplication problems with their inverse partitive and quotitive division problems, and (ii) connecting the invariant structures between these problem structures across natural numbers and proper/improper fractions. The instructional approach for initiating making these connections is based on the *Bianshi* approach, with its design principle of problem variation (Sun, 2019). The variation principle is used for designing and analyzing textbooks (Sun, 2019) and for teacher professional development (Han et al., 2017). However, as far as we know, there are no empirical studies providing in-depth insights into students' work. So, we ask: *How do students' solving modelling pathways change during the work with varied problem sets?*

2023. 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. 387-494). PME 46.

THEORETICAL FRAMEWORK

Variation principle used for the design of the varied problem sets

Considering the two connections mentioned above, we used three types of variations (Sun, 2019): (i) One problem multiple changes (OPMC): Initially presenting and solving a problem, and later solving variations of the initial problem (by varying condition/s). (ii) One problem multiple solutions (OPMS): Providing the opportunity to solve a problem using different strategies. (iii) Multiple problems one solution (MPOS): Using the same strategy to solve a set of identical structure problems.

For this, five varied problem sets were designed (Table 1; Zorrilla et al., 2022). For focusing students' attention on the inverse relationship, we used the variation OPMC for composing each problem set of three problems with the same situation but different unknowns: A multiplication problem (M), e.g., "We have [number of groups] packages. Each package contains [quantity per group] kilos. How many kilos do we have in total?"; a partitive division problem (P), e.g., "We have [number of groups] packages. All the packages contain the same number of kilos. We have [total quantity] kilos in total. How many kilos does each package have?", and a quotitive division problem (Q): "We have some packages. Each package contains [quantity per group] kilos. We have [total quantity] kilos in total. How many packages do we have?".

Table 1: Five problem sets and extra problems designed

	Set 1	**Set 2**	**Set 3**	**Set 4**	**Set 5**
Situation	3 packages of 4kg are 12kg	8 packages of $\frac{1}{2}$kg are 4kg	$5\frac{1}{2}$ packages of 2kg are 11kg	8 packages of $\frac{1}{10}$kg are $\frac{4}{5}$kg	$2\frac{1}{4}$ packages of 2kg are $4\frac{1}{2}$kg
M	**1.1M** 3×4	**2.3M** $8 \times \frac{1}{2}$	**3.2M** $5\frac{1}{2} \times 2$	**4.1M** $8 \times \frac{1}{10}$	**5.1M** $2\frac{1}{4} \times 2$
P	**1.2P** $12 : 3$	**2.1P** $4 : 8$	**3.3P** $11 : 5\frac{1}{2}$	**4.3P** $\frac{4}{5} : 8$	**5.2P** $4\frac{1}{2} : 2\frac{1}{4}$
Q	**1.3Q** $12 \div 4$	**2.2Q** $4 \div \frac{1}{2}$	**3.1Q** $11 \div 2$	**4.2Q** $\frac{4}{5} \div \frac{1}{10}$	**5.3Q** $4\frac{1}{2} \div 2$
	Extra problems (other situations and numbers)				
M	**E1.4M** 8×2	**E2.6M** $10 \times \frac{2}{5}$			**E5.4M** $5\frac{1}{4} \times 3$
P	**E1.5P** $20 : 5$	**E2.4P** $2 : 20$			**E5.5P** $8\frac{4}{5} : 2\frac{1}{5}$
Q	**E1.6Q** $24 \div 4$	**E2.5Q** $2 \div \frac{1}{4}$			**E5.6Q** $7\frac{1}{2} \div 2$

The variation OPMC was applied across the five problem sets to maintain the structure of the problem (same unknown), while changing the number types involved (Table 1). The variation MPOS aimed at inviting students to use the same strategy in new number types. Furthermore, students worked on the problems in small groups, and then the teacher shared and discussed different strategies in the whole class (working in a format OPMS). The order of problems varied, e.g., 1.1M–1.2P–1.3Q, but 2.1P–2.2Q–2.3M.

To support students' "disregard" of the context and to initiate that they switch to generalizing independently of the given context, some extra problems (Table 1) were

included (with other context situations and other numbers). This variation aimed at encouraging students to extrapolate the discovered structures to other contexts and, thus, generalize the structure (Marton, 2015).

Framework to investigate students' pathways

According to classical *modelling* cycles, students can work on a word problem (textual representation) by *structuring* the underlying situation of the problem (contextual representation; graphical or mental representation), which would allow them to *mathematize* by choosing the operation that solves the problem (symbolic representation) to, then, proceed to *calculate* the result (numerical representation). The student would then be expected to *interpret* the operation and result according to the structure of the situation and *validate* the result in that situation (vom Hofe & Blum, 2016). Typical obstacles in completing the modelling cycle can be related to inadequate mental models (understood as the translation knowledge given by standard interpretations of operations, see Fischbein et al., 1985; Greer, 1992), poor reading comprehension that can lead to a deviant mental representation (Dröse & Prediger, 2020), or superficial interpreting and validating processes that can hinder students from recognizing their errors (Verschaffel et al., 2000).

In this study, we adapted the modelling cycle to capture also *emergent modelling* (see Gravemeijer, 2007), i.e., the processes involved in extending the meaning of operations from natural numbers to fractions and draw upon the inverse relationship between division and multiplication. Emergent modelling partially draws upon the same processes as modelling, but instead of choosing between the existing mental models, students are encouraged to draw upon informal strategies (Empson & Levi, 2011) and connect them to the newly emerging mental models or to draw upon inverse relationships among multiplication and division problems by discovering the commonality and the inverse relationship in the problem texts.

METHOD

Methods of data gathering. This paper is part of a larger study in which we have examined students' modelling pathways when they solved the problem sets designed. 17 Spanish sixth graders (11-12 years old) solved the problem sets in small groups. The data corpus comprises students' small-group discussions (audio-recorded and transcribed) and students' individual worksheets.

Methods of qualitative data analysis. In order to capture students' modelling pathways in the audio-recorded data, we applied a combined deductive and inductive qualitative procedure (Zorrilla et al., submitted). First, each student utterance was deductively coded regarding the addressed representation (textual, graphical or contextual, symbolic, and numerical representation) and whether students referred to other problems. Second, each student utterance was segmented into sense-making units referring to the same representation/s or idea/s and we performed an inductive coding of (emergent) sub-processes (e.g., related to the process of *structuring*, the sub-processes of *contextualizing* the data of the problem and *reporting* the mental

contextual representation and/or graphical representation). From this deductive and inductive coding, we obtained 408 chains (24 problems and 17 students) of sub-processes with the representation addressed. These chains of sub-processes were systematically compared and classified into reoccurring patterns. In case that the information of the chain was not enough, this chain was classified as *non-communication*. This analysis revealed five typical students' individual modelling pathways that partially deviate from the intended modelling trajectories (Zorrilla et al., submitted): (i) Direct Translation pathway: Students directly translate the problem text into symbolic and numerical representations. (ii) Contextualizing the Chosen Operation pathway: Students directly articulate the chosen operation in connection with mental contextual representation of the problem text that allows them to contextualize the operation. (iii) Informal Strategies pathway: Students solve the problem using a graphical or contextual representation, and they may (or may not) mathematize using the symbolic representation but without reaching the formal operation. (iv) Operational Connection pathway: Students find or validate the result by making connections to other problems but not by relating mathematical problem structures. (v) Relational Connection pathway: Students connect problems using the inverse relationship between multiplication and division.

Analytic step for the current paper. Here, we extend this previous work by examining the changes in students' modelling pathways along the varied problem sets. For this longitudinal analysis, we focus on Lola, Tíscar, Jose, Luis, Ana, and Alba. The changes indicate how students are learning to solve multiplicative structure problems in the problem sets designed with the variation principle.

RESULTS

Most students used the *Direct Translation pathway* in the first multiplication problem (Set 1: 1.1M: "We have 3 packages. Each package contains 4 kilos. How many kilos do we have in total?"). In fact, all six students directly translated the multiplication problem text 1.1M successfully into symbolic and numerical representation. The excerpt of transcript exemplifies this *Direct Translation pathway* in which the small group of Lola, Tíscar, and Jose focused on comparing their operations and result:

Lola: [*After writing down 3 × 4 = 12*] 12 kilos… I know. It's 12 kilos.

Jose: It's 12 kilos.

Tíscar: Well, yes… 4 times 3 would be 12. Then that's it.

Also, Luis, Ana and Alba, started by individually reading the problem, then they proceeded to mathematize it by immediately choosing the multiplication operation and calculated the result to solve problem 1.1M (*Direct Translation pathway*).

In contrast to their first encounter with the problem sets that was based on natural numbers, the students' pathways then changed from a strong focus on the *Direct Translation pathway* to a much broader variety of pathways. In the varied problem sets designed, we used the variation OPMC to focus students' attention on the inverse

relationships between problems. Each problem set was constructed by three problems with the same situation and numbers, but different unknowns. This variation principle seems to help students to make deeper connections between problems based on the inverse relation (*Relational Connection pathway*). This was the case for Tíscar and Jose: After solving the quotitive and partitive division problems 2.1Q and 2.2P in Set 2, they solved the inverse multiplication problem 2.3M ("We have 8 packages. Each package contains ½ kilos. How many kilos do we have in total?") by a *Relational Connection pathway*, saying:

Tíscar: It's like the other way around. It's the inverse.

Jose: They are inverse operations [*both had written 0.5 × 8 = 4.0*].

However, some students developed superficial connections based on the numbers instead of the mathematical structure of the problem (*Operational Connection pathway*) and used these superficial connections either for obtaining or validating the unknown respectively the operation used to calculate the unknown. This was the case, e.g., for Lola who solved the partitive division problem of Set 1 (1.2P: "We have 3 packages. All the packages contain the same number of kilos. We have 12 kilos in total. How many kilos does each package have?") after the multiplication problem 1.1M with the same numbers. In her dialogue with Tíscar, we observe that Lola agreed that "4 kilos" is the result since it matches with the numbers of the multiplication problem that she had solved before. Both students obtained the results from the other problem and tried to search for a suitable operation. For Lola, this process resulted, first, in an incorrect operation providing the operation 3 : 12.

Tíscar: It's saying it here: "Each package contains 4 kilos" [*referring to multiplication problem 1.1M*].

Lola: […] Yes, because the first problem [*referring to multiplication problem 1.1M*] says it.

Tíscar: […] [*After Lola had written 3 : 12 = 4*] Lola, I think it's the other way around. Because if you do 3 divided by 12, the result is different.

Lola: So… 12 divided by 3?

In some cases, these superficial connections were used to validate the operations used to obtain the result. This is the case, e.g., for Ana in the multiplication problem in Set 2 (2.3M: "We have 8 packages. Each package contains ½ kilos. How many kilos do we have in total?"). Following a *Contextualizing the Chosen Operation pathway*, she structured the situation by contextualizing that there are 8 packages with ½ kilo each, but she realized that her operation (8 ÷ 0.5) was not correct since the result should have been 4 (looking at 2.2Q previously solved) (*Operational Connection*).

Ana: 8 divided by 0.5… Because we have 8 packages, and each package weighs half a kilo… "How many kilos do we have in total?".

Ángela: […] 4 would have to be the result…

Ana: Yes, 4 is the result. I just got it wrong.

Teacher: And why are you sure that 4 is the result?

Ana: Because it says it here [*pointing at the prevision quotitive division problem 2.2Q in the same problem set*]: "We have 4 kilos in total".

We used the variation OPMC across the five problem sets maintaining the structure of the problem (same unknown) and changing the numbers involved (from natural numbers to fractions). Supported by this variation, some students changed from *Direct Translation pathway* (with natural numbers) to *Contextualizing the Chosen Operation pathway* (with fractions). This change shows the need for students to explicitly construct a contextual representation of the operation with fractions. This is the case for Luis, who initiated the solution process of the multiplication problem 2.3M by structuring the problem text. This allowed him, after transforming from the fraction representation to the decimal number representation, to mathematize the problem text using the formal operation: "0.5 multiplied by 8 packages".

Luis: It [*problem text*] says that each package weighs… 1 over 2, which is equal to 0.5 kilograms. So, each package weighs this and it says "we have 8 packages" … So, 0.5 multiplied by 8 packages.

This variation (from natural numbers in Set 1 to fractions in Set 2) in format OPMS (with discussion of different strategies to solve the problems in the whole class), helped students to use informal strategies (they changed from *Direct Translation pathway* with natural numbers to *Informal Strategies pathway* with fractions). The following excerpt of the dialogue between the teacher and Alba shows the use of an informal strategy in the quotitive division problem of Set 5 (5.3Q: "We have some packages. Each package contains 2 kilos. We have 4 ½ kilos in total. How many packages do we have?").

Alba: As 4.5 are the total of kilos we have, I've divided the kilos by each package.

Teacher: So, you've been drawing the packages…

Alba: Yes, because each package has 2 kilos. So, I've made the packages…

Teacher: OK. You've done two equal packages but the third one is different. Why?

Alba: Because we don't have 6 kilos.

Teacher: […] And how many kilos there are in the third one?

Alba: 0.5.

Teacher: So, how many packages do we have?

Alba: We have 2.25.

DISCUSSION AND CONCLUSIONS

We have analyzed the changes in students' pathways through the problem sets designed to identify characteristics of students' learning. These characteristics give us relevant information with regard to how students are learning to solve multiplicative structure word problems (in the challenging shift from natural numbers to fractions,

Greer, 1992), when the variation principle (Marton, 2015; Sun, 2019) is used in our design to help students overcome the difficulties documented in the literature.

The design helped some students to draw *Relational Connections* between multiplicative structure problems. This was observed, e.g., when Tíscar and Jose drew on the inverse relationship to choose the operation of a multiplication problem by relating it to the partitive and quotitive problems solved previously. For other students, the deliberate variation of problems helped them to move from natural numbers to fractions, pushing them to use instead of the most common pathway used with natural numbers (*Direct Translation pathway*), the *Contextualizing the Chosen Operation pathway* or the *Informal Strategies pathway* (with fractions). The change from *Direct Translation* to *Contextualizing the Chosen Operation* seems to indicate that students have a contextual mental representation of the problem (although not verbalizing it when working with natural numbers). The introduction of fractions seems to push them to make explicitly this contextual mental representation showing the ability to control the generalized use of certain models across the various mental processes (vom Hofe & Blum, 2016). The change from *Direct Translation* to *Informal Strategies pathway* shows how students could give meaning to operations with fractions through informal strategies (Empson & Levi, 2011) before mathematizing with the formal operation.

Considering these results, we observe the powerful role that OPMC played on engaging students in the two types of connections (Yanhui, 2018) to overcome well-documented difficulties (Greer, 1992). The variation principle enhanced some students drawing connections and helped them in the transition from natural numbers to fractions.

However, there were students such as Lola that developed superficial connections based on the numbers instead of the mathematical structure of the problem (*Operational Connection pathway*), which hindered them from constructing meanings for multiplication or division of fractions. This observation is in line with Dröse and Prediger's (2020) findings in a completely different context, some students reflected on the role that variation had on the structure of the new problem, while others stayed on more superficial connections (*Relational Connection pathway* vs. *Operational Connection pathway* in our study). Nevertheless, these superficial connections appeared to be already useful in some cases at least for the mental process of validation, since some students followed this pathway to validate the operation and result obtained.

ACKNOWLEDGEMENTS

This research was supported by Ministerio de Universidades (Spain): FPU grant FPU19/02965 and EST21/00333.

REFERENCES

Dröse, J., & Prediger, S. (2020). Enhancing fifth graders' awareness of syntactic features in mathematical word problems: A Design Research study on the variation principle. *Journal für Mathematik-Didaktik*, *41*(2), 391-422. https://doi.org/10.1007/s13138-019-00153-z

Empson, S. B., & Levi, L. (2011). *Extending children's mathematics: Fractions and decimals*. Heinemann.

Fischbein, E., Deri, M., Nello, M. S., & Marino, M. S. (1985). The role of implicit models in solving verbal problems in multiplication and division. *Journal for Research in Mathematics Education*, *16*(1), 3-17. https://doi.org/10.2307/748969

Gravemeijer, K. (2007). Emergent modelling as a precursor to mathematical modelling. In W. Blum, P. L. Galbraith, H.-W. Henn, & M. Niss (Eds.), *Modelling and applications in mathematics education: The 14th ICMI Study* (pp. 137-144). Springer.

Greer, B. (1992). Multiplication and division as models of situation. In D. Grows (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 276-295). McMillan.

Han, X., Gong, Z., & Huang, R. (2017). Teaching and learning mathematics through variation in Lesson Study. In R. Huang & Y. Li (Eds.), *Teaching and Learning Mathematics through Variation* (pp. 267-293). Sense Publishers.

Marton, F. (2015). *Necessary conditions of learning*. Routledge.

Sun, X. H. (2019). Bridging whole numbers and fractions: problem variations in Chinese Mathematics textbook examples. *ZDM – Mathematics Education*, *51*(1), 109-123. https://doi.org/10.1007/s11858-018-01013-9

Verschaffel, L., Greer, B., & De Corte, E. (2000). *Making sense of word problems*. Swets & Zeitlinger.

Vom Hofe, R., & Blum, W. (2016). "Grundvorstellungen" as a Category of Subject-Matter Didactics. *Journal für Mathematik-Didaktik*, *37*, 225-254. https://doi.org/10.1007/s13138-016-0107-3

Yanhui, X. (2018). On "One Problem Multiple Change" in Chinese "Bianshi" mathematics teaching. *The Teaching of Mathematics*, *21*(2), 80-96.

Zorrilla, C., Ivars, P., Fernández, C., & Llinares, S. (2022). Solving multiplicative structure problems with fractions in Primary School: Prior strategies to a teaching experiment. In J. Hodgen, E. Geraniou, G. Bolondi, & F. Ferreti (Eds.), *Proc. 12th Congress of the European Society for Research in Mathematics Education (CERME12)*. Free University of Bozen-Bolzano and ERME.

Zorrilla, C., Roos, A.-K., Fernández, C., Llinares, S., & Prediger, S. (submitted). Connecting operation-choice problems by the variation principle: Sixth graders' operational or deeper relational pathways. *Submitted journal manuscript*.

אוניברסיטת חיפה
University of Haifa
جامعة حيفا

הפקולטה לחינוך كلية التربية
THE FACULTY OF EDUCATION

9 789659 311248