THE IMPACT OF STUDENT AND TASK CHARACTERISTICS ON VALIDATION PROCESSES IN DATA-BASED MODELLING

Timo Kosiol[1], Marielena Menzel[2], Sebastian Geisler[2], Stefanie Rach[1]

[1]OVGU Magdeburg, [2]University of Potsdam

Validating is considered an important process for successful modelling. It includes critically evaluating and, if necessary, revising (intermediate) models or results during a modelling process. In an experimental study with N=561 upper secondary students, we compare validation processes for students modelling the decay of beer foam using real data (either self-collected or provided) with those for students modelling smoothed data. We examine whether student characteristics (motivation and content knowledge) or task characteristics (the type of data) predict validation processes. Results indicate that a model's critical evaluation is predicted by content knowledge, while the formulation of suggestions for improvement is predicted by the type of data. We discuss the role of student and task characteristics to stimulate validation processes.

INTRODUCTION

Modelling competence—a central goal of mathematics education—refers to the ability to solve real-world problems using mathematical models(Cevikbas et al., 2022). Validation processes are considered essential for successful modelling, as students evaluate their potentially intermediate models or results against prior models or results during modelling processes. However, validation has been found to place significant demands on students, and student and task characteristics might influence whether they engage in such validation processes Thus, we investigate how student characteristics, such as motivation and knowledge, and task characteristics, such as the type of data, impact students' engagement in validation processes. In addition, we describe various validation processes observed when students model data using functions.

THEORETICAL BACKGROUND

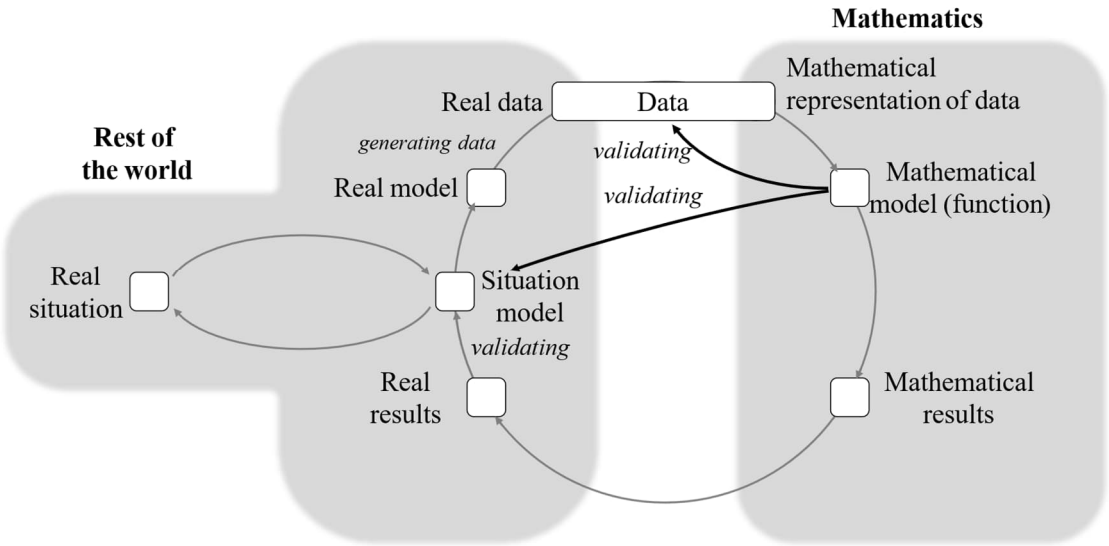
Modelling refers to the process of solving real-world problems using mathematical models (Niss & Blum, 2020). In the past years, data-based modelling has received increasing attention and describes the process of setting up a mathematical model based on data (and information about a phenomenon) to better understand a real-world phenomenon from which the data arose and make predictions regarding the future (e.g., Mohr & Ufer, 2023). Modelling processes are described using modelling cycles (see Figure 1): Students construct a mental model of the situation (situation model) and simplify and structure it (real model). In data-based modelling, students draw on data that can be collected, for example, through experiments (Jahn et al., 2022). This data can be represented mathematically, e.g., in a coordinate system, and described by a mathematical function (Engel, 2009). Mathematization can be grounded in the data as

2026. In H. Krzywacki, M. Hähkiöniemi, E. Haataja, M.S. Hannula & P. Hästö (Eds.). *Proceedings of the 49th Conference of the International Group for the Psychology of Mathematics Education: Research Reports*, pp. 859-866. PME.

well as assumptions about the real-world phenomenon (English & Watson, 2018; Stillman & Brown, 2021). The function can be used to obtain mathematical results. After interpreting these results, validating them based on the students' situation model is described as an important and challenging sub-process of modelling (Blum, 2015).

Validating: Qualitative analyses of modelling processes show that validation is not only based on the fit between real results and the situation model, but also occurs during modelling, considering the fit between potentially intermediate models or results, for example, the mathematical model and the situation model (Czocher, 2018). In Figure 1, possible validation processes regarding the mathematical model, on which we focus, are represented using arrows. Validating encompasses processes in which (i) the fit is critically evaluated and, if necessary, (ii) individual results or parts of the model are revised (Cevikbas et al., 2022; Hankeln et al., 2019).

Figure 1: An adjusted modelling cycle for data-based modelling
(cf. Blum & Leiss, 2007; Eichler & Vogel, 2013)



Task characteristics, such as the type of data to be modelled, may influence validation processes. Modelling real data is hypothesized to stimulate validation processes when the variation within the data is evaluated regarding how well a function fits the data (Engel, 2009). In contrast, when modelling smoothed data—that is fake data that fit the intended function perfectly—validating the mathematical model becomes meaningless once the intended function has been identified. Hands-on experiments have been proposed as an effective method for collecting real (empirical) data for modelling processes and as a means of validating models and results during modelling (Jahn et al., 2022; Zell & Beckmann, 2009). Collecting real data through experimentation may improve students' understanding of the context compared to simply providing them with data (Carreira et al., 2020). However, there is a lack of experimental studies to test the influence of the type of data on validation processes.

Student characteristics: Successful validation processes require students to recognize errors in their results or models and respond appropriately to these errors (Goos, 2002).

Therefore, it is plausible that prior content knowledge influences validation processes (see Mohr & Ufer, 2023 for the whole modelling processes). By content knowledge, we refer to conceptual and procedural knowledge about functions. In addition to content knowledge, motivational characteristics such as self-beliefs and value beliefs may influence modelling processes (Schukajlow et al., 2022). In expectancy-value theory, expectancies of success and task values are posited to influence both the choice to engage in a task and performance (Eccles & Wigfield, 2020). Expectancies of success are a construct similar to self-concept, referring to students' beliefs about their mathematical abilities. Task values encompass intrinsic value (i.e., anticipated enjoyment while doing the task), utility value (i.e., relevance attributed to a task regarding daily life, career, or upcoming events), attainment value (i.e., perceived importance of a task), and costs (i.e., anticipated negative consequences when doing the task). The inclusion of costs within task values or as a separate construct has been debated (Barron & Hulleman, 2015). While the relationship between modelling performance and motivation has been investigated previously (e.g., Schukajlow & Krug, 2014), we focus on the relation between validation processes and motivational characteristics in combination with content knowledge.

RESEARCH QUESTIONS (RQ)

Focussing on task and student characteristics in validation processes when students model data with functions, we address the following research questions:

1. To what extent do the type of data (self-collected real data, provided real data, provided smoothed data) or student characteristics (content knowledge, self-concept, task value, costs) predict whether students critically evaluate their model?
2. To what extent do the type of data and student characteristics predict whether students propose a suggestion for improvement of the model?
3. Which suggestions for improvement do students propose?

METHOD

Sample and design

The study is part of the project *Experiments to foster modelling competences and motivation in mathematics* (Ex2MoMa; Jonscher et al., 2024). We sampled $N = 561$ students from 27 classes (10th and 11th grade) in Germany (52% female, $M_{age} = 16$ years). Initially, student characteristics were assessed. Subsequently, students worked on three modelling tasks across three consecutive mathematics lessons, integrated into regular teaching and directed by trained research assistants. Students worked in pairs in a self-regulated learning environment, supported by a guiding manual and answering the task individually.

Each task required modelling a real-world situation based on data using either a linear or exponential function. For all lessons, classes were randomly assigned to one of two experimental conditions or the control group: Students in the first experimental

condition performed a hands-on experiment first and subsequently modelled the self-collected data (EG1, $N = 215$), while the second condition modelled given real data (EG2, $N = 160$). The control group modelled smoothed data (CG, $N = 186$).

In the second lesson, students modelled the decay of beer foam, which can be mathematized using an exponential function. Due to students' unfamiliarity with least-squares techniques, they were expected to mathematize the data using at least one pair of data points and were provided with a GeoGebra applet with restricted functionality to graphically optimize the function. Students were explicitly asked to check their mathematical models and, if necessary, to provide suggestions for improvement.

Instruments

Content-specific prior knowledge of linear and exponential functions was measured with a 7-items test (see Figure 2). Items demand students to identify parameters of a linear or exponential function, given a table or graph, or to match a graph to a given equation (see Jonscher et al., 2024). We scaled the test using an item-response-theory model with good item fit and acceptable WLE reliability (*WLE* = .62).

Figure 2: Item example for the content knowledge test on functions

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

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

$f(x) =$ _____

Motivational student characteristics were assessed with three scales: Mathematical self-concept (4 items, $\alpha = .94$, example item: "*In mathematics, I am good.*"), task values focusing on utility, attainment, and intrinsic value (19 items, $\alpha = .94$, example item: "*I find mathematics exciting.*"), and costs (6 items, $\alpha = .88$, example item: "*Learning math exhausts me.*"). All items were answered on a six-point Likert-scale (1 = "not at all true", 6 = "totally true"). Each scale was assessed with good reliability.

Validation processes were coded from students' individual written responses within the learning environment. Two dichotomous codes were assigned to indicate whether (i) the students *critically evaluated* their model as *not fitting sufficiently well* to the function or situation, as well as (ii) whether a specific *suggestion for improvement* was made. Interrater objectivity of 15% double-coded answers was satisfactory (critical evaluation: Cohen's $\kappa = .83$; suggestion for improvement: Cohen's $\kappa = .95$).

Data analysis

Missing data (range: 0%-17.8%, mean: 9.4%) resulted primarily due to missing at one data collection. However, Little's test of missing completely at random was significant ($\chi^2(53) = 103.9$, $p < .001$). Assuming data to be missing at random, we used multiple imputation (with *mice* in R) to generate 20 sets of imputed values. Addressing RQ 1

and 2, the data were analysed using two generalized linear models to predict whether students (i) would evaluate their mathematical model as not fitting sufficiently well and (ii) would formulate a suggestion for improvement by the experimental condition (type of data) and student characteristics. In both models, we controlled for model quality, in the second model, the outcome of the first, the variable indicating whether students critically evaluate one's own model, was additionally included as a control variable. Model quality was estimated using the difference between the residual sum of squares (RSS) of the function developed by the students and the RSS of the best-fitting function determined by the least-square method. To account for skewness, we have logarithmized the difference, and a lower value indicates a better fit of the function to the data. To answer RQ 3, we categorized the students' suggestions for improvement inductively, and present typical answers.

RESULTS

RQ1: Critical evaluation of one's own model as not good enough

Results of the two generalized linear models are presented in Table 1. Controlling for model quality, using self-collected (Estimate = 0.56, $p = .073$) or provided (Estimate = 0.34, $p = .311$) real data does not significantly affect the critical evaluation of the model. The low intercept in Model 1 indicates that a student of the control group with average characteristics and a good model fit has a probability of 11.7% to critically evaluate the own model as not good enough. Descriptively, the probability for a critical evaluation increases in both conditions. Regarding student characteristics, students with higher content knowledge critically evaluate their model as not good enough than the other students (Estimate = 0.28, $p = .041$). While self-concept does not appear to relate to model evaluation, descriptively, students with higher task values (Estimate = -0.23, $p = .114$) and costs (Estimate = -0.21, $p = .257$) appear less likely to critically evaluate their own models.

RQ2: Formulation of a suggestion for improvement

As expected, students who critically evaluate their model as not good enough are more likely to formulate suggestions for improvement (Estimate = 1.22, $p < .001$, Table 1). Compared with modelling smoothed data, students who worked with real data were more likely to formulate suggestions for improvement, regardless of whether this data was collected themselves (Estimate = 0.63, $p = .019$) or provided (Estimate = 0.55, $p = .061$). While students in the control group have a probability of 55% to formulate a suggestion for improvement, this probability increases to 70% (EG1) or 68% (EG2) in the experimental conditions. Regarding student characteristics, only students' task values are a significant factor (Estimate = 0.28, $p = .042$).

RQ3: Types of suggestions for improvement

Students formulate various suggestions for improvement. Students generally suggested to collect data more carefully (e.g., "Avoid measurement error when collecting data") or specifically suggested changes in the experiment (e.g., "Always measure the beer

foam at the edge"; "Collect more data."). Suggestions focused on the function include replacing the type of mathematical function (e.g., using an exponential instead of a linear function) or modifying its parameters to better fit the data. Rarely, students suggested calculating the decay factor using multiple pairs of values to represent the entire dataset. One suggestion that occurred most frequently in the control group (using smoothed data) was to round calculated values less to get a more accurate model.

Table 1: Results of generalized linear models predicting validation processes

Outcome	**Critical model evaluation**	**Suggestion for improvement**
Predictors	Estimate (*SE*)	Estimate (*SE*)
Intercept	-2.15 (0.25)***	0.20 (0.17)
Model quality	0.32 (0.07)***	-0.04 (0.05)
Critical model evaluation		1.22 (0.29)***
EG1	0.56 (0.31)^{t}	0.63 (0.27)*
EG2	0.34 (0.34)	0.55 (0.29)^{t}
Content knowledge	0.28 (0.14)*	0.15 (0.13)
Task value	-0.23 (0.15)	0.28 (0.14)*
Self-concept	0.01 (0.20)	0.17 (0.18)
Costs	-0.21 (0.18)	0.22 (0.16)

Note. *SE* = Standard error; $^{*}p<.05$, $^{**}p<.01$, $^{***}p<.001$, $^{t}p<.10$. Student characteristics have been z-standardized; EG1: Self-collected real data, EG2: Provided real data; a lower value in *model quality* indicates a better fit, 0 the best possible fit.

DISCUSSION

This experimental study investigated the impact of the type of data as well as students' knowledge and motivational characteristics on whether students engage in validation processes. Our findings indicate that modelling real data may stimulate validation processes compared to smoothed data (Zell & Beckmann, 2009). More specifically, our results suggest that it may help students generate suggestions for improvement. Regarding experimenting, there doesn't seem to be a benefit when students collect the data themselves. This may be due to a limitation of our study; we focused on the existence but not the quality of validation processes. As hands-on experiments have been theorized to help students understand the context of real phenomena (Carreira et al., 2020), future studies also include the quality of validation processes.

With regard to student characteristics, our study extends the findings of Schukajlow and Krug (2014) concerning whole modelling processes to validation processes. In line with qualitative results by Goos (2002), we have found that students rely on content knowledge to determine whether a model fits the data sufficiently well. We cannot

relate content knowledge to formulation suggestions for improvement; however, this may be due to the focus on the mere existence of suggestions rather than their quality. This is also consistent with findings by Goos (2002), who found that students may react inadequately to errors or even recognize errors that do not exist. Future studies should investigate the relevance of content knowledge for the quality of the suggestions. In contrast to content knowledge, students' task value is significantly related to the formulation of suggestions for improvement. Students with high value beliefs regarding mathematics are more likely to formulate suggestions. This finding aligns with theories proposing that students with a high motivation engage more deeply with modelling problems (Schukajlow et al., 2022).

Beyond merely noting the existence of validation processes, our study is also limited because students did not always make their validation processes explicit in the manual. Although explicitly prompting to validation processes, the written answers do not always allow to deduce whether and which validation processes occurred.

Future studies may investigate the quality of validation processes as well as the extent to which these processes influence the overall performance of the entire modelling process. Summarizing, our results highlight that task and student characteristics affect validation processes. However, the pattern of impact is complex: High content knowledge supports the critical evaluation of the model, while motivational characteristics and the use of real data appear to support the revision of the model.

Funding: This study is funded by the German federal ministry of education and research under grant No. 16MF1028A.

References

Barron, K. E. & Hulleman, C. S. (2015). Expectancy-value-cost model of motivation. In J. D. Wright (Ed.), *Int. enc. of the social & beh. sciences* (2nd ed., Vol. 8, pp. 503-509). Elsevier.

Blum, W. (2015). Quality teaching of mathematical modelling: What do we know, what can we do? In S. J. Cho (Ed.), *The Proceedings of the 12th International Congress on Mathematical Education* (pp. 73-96). Springer.

Blum, W. & Leiss, D. (2007). How do students and teachers deal with modelling problems? In C. Haines, P. Galbraith, W. Blum, & S. Khan (Eds.), *Mathematical Modelling* (pp. 222-231). Woodhead Publishing. https://doi.org/10.1533/9780857099419.5.221

Carreira, S., Baioa, A. M., & Werle de Almeida, L. M. (2020). Mathematical models and meanings by school and university students in a modelling task. *Avances de Investigación en Educación Matemática* 17, 67-83. https://doi.org/10.35763/aiem.v0i17.308

Cevikbas, M., Kaiser, G., & Schukajlow, S. (2022). A systematic literature review of the current discussion on mathematical modelling competencies: State-of-the-art developments in conceptualizing, measuring, and fostering. *Educational Studies in Mathematics*, *109*(2), 205-236. https://doi.org/10.1007/s10649-021-10104-6

Czocher, J. A. (2018). How does validating activity contribute to the modeling process? *Educational Studies in Mathematics*, *99*(2), 137-159.

Eccles, J. S. & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology*, *61*, Article 101859.

Eichler, A. & Vogel, M. (2013). *Leitidee Daten und Zufall* [Key concept: Data and chance] (2 ed.). Springer. https://doi.org/10.1007/978-3-658-00118-6

Engel, J. (2009). *Anwendungsorientierte Mathematik: Von Daten zur Funktion* [Application-oriented mathematics: From data to mathematical functions]. Springer.

English, L. D. & Watson, J. (2018). Modelling with authentic data in sixth grade. *ZDM*, *50*(1-2), 103-115. https://doi.org/10.1007/s11858-017-0896-y

Goos, M. (2002). Understanding metacognitive failure. *The Journal of Mathematical Behavior*, *21*(3), 283-302. https://doi.org/10.1016/S0732-3123(02)00130-X

Hankeln, C., Adamek, C., & Greefrath, G. (2019). Assessing sub-competencies of mathematical modelling—development of a new test instrument. In G. A. Stillman & J. P. Brown (Eds.), *Lines of Inquiry in Mathematical Modelling Research in Education* (pp. 143-160). Springer. https://doi.org/10.1007/978-3-030-14931-4_8

Jahn, B., Friedrich, S., Behnke, J., Engel, J., Garczarek, U., Munnich, R., Pauly, M., Wilhelm, A., Wolkenhauer, O., Zwick, M., Siebert, U., & Friede, T. (2022). On the role of data, statistics and decisions in a pandemic. *Advances in Statistical Analysis*, *106*(3), 349-382.

Jonscher, M., Menzel, M., Geisler, S., & Rach, S. (2024). Modelling performance using functions: Relation to person characteristics and different solution approaches. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.), *Proc. 47th Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 3, pp. 113-120). PME.

Mohr, M. & Ufer, S. (2023). Measuring data-based modeling skills in a collaborative setting. In M. Ayalon, B. Koichu, R. Leikin, L. Rubel, & M. Tabach (Eds.), *Proc. 46th Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 3, pp. 355-362). PME.

Niss, M. & Blum, W. (2020). *The learning and teaching of mathematical modelling*. Routledge. https://doi.org/10.4324/9781315189314

Schukajlow, S., Krawitz, J., & Carreira, S. (2022). The influence and role of affective aspects in mathematical modelling. In C. Fernández, S. Llinares, A. Gutiérrez, & N. Planas (Eds.), *Proc. 45th Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 1, pp. 158-161). PME.

Schukajlow, S. & Krug, A. (2014). Are interest and enjoyment important for students' performance? In C. Nicol, S. Oesterle, P. Liljedahl, & D. Allan (Eds.), *Proceedings of the joint meeting of PME 38 and PME-NA 36* (Vol. 5, pp. 129.136). PME.

Stillman, G. A. & Brown, J. P. (2021). Modeling the phenomenon versus modeling the data set. *Mathematical Thinking and Learning*, *25*(3), 270-295.

Zell, S. & Beckmann, A. (2009). Modelling activities while doing experiments to discover the concept of variable. In V. Durand-Guerrier, S. Soury-Lavergne, & F. Arzarello (Eds.), *Proceedings of CERME 6* (pp. 2216-2225). INRP.