



HAL
open science

Students' approaches concerning model validation and model improvement when solving modelling tasks with experiments

Sebastian Geisler

► To cite this version:

Sebastian Geisler. Students' approaches concerning model validation and model improvement when solving modelling tasks with experiments. Twelfth Congress of the European Society for Research in Mathematics Education (CERME12), Feb 2022, Bozen-Bolzano, Italy. hal-03759004

HAL Id: hal-03759004

<https://hal.science/hal-03759004>

Submitted on 23 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Students' approaches concerning model validation and model improvement when solving modelling tasks with experiments

Sebastian Geisler

University of Hildesheim, Germany; geisler@imai.uni-hildesheim.de

Mathematical modelling is important yet challenging. Especially the validation step is a hurdle for many students. One approach, discussed in the literature, for fostering students' validation competence is to use modelling tasks that use data gathered in experiments. Analyzing the validation of 71 students in two modelling tasks with experiments, we found that many students struggle with validation. In particular, most students seem to put more trust in their model than in their experimental data and therefore suggest to improve experiments instead of working on an alternative model. Stimulated recall interviews show that students' lack of mathematical knowledge as well as a lack of understanding of the relation between data and model can be the reasons for these findings.

Keywords: Experiments, mathematical modelling, model-validation.

Introduction

Mathematical modelling is a central mathematical competence, which is reflected in several curricular documents around the world (e.g., National Governors Association Center for Best Practices and Council of Chief State School Officers, 2010). Nevertheless, modelling remains complex and challenging for students. Previous studies have shown that especially the validation of one's models is a hurdle in the modelling process for many students (e.g. Blum & Leiß, 2007). Therefore, supporting students when validating their models is necessary. According to Engel (2010) real data is needed for authentic modelling and validation, since real data usually does not fit perfectly to an intended model. In contrast, using already smoothed or unrealistic data that fits well to an intended model, leads to a modelling process in which the relevance of validation is not stressed (Engel, 2010). One approach to integrate real data into the modelling process is to use modelling tasks with experiments in which students first perform experiments and gather data that can be used for the subsequent modelling (cf. Zell & Beckmann, 2009). This approach is frequently discussed in practical teaching literature (e.g. Ludwig & Oldenburg, 2007) but little empirical evidence for the anticipated positive effects on students' validation competence exists.

The project *Mathematical Modelling with Experiments* (MaMEx) follows the objectives to design and evaluate modelling tasks with experiments and to gather insights into students' validation processes. In a pilot study, Geisler (2021) used a modelling task with experiment concerning the decay of beer-froth and analyzed students' validation approaches. Students in this pilot study still struggled with validation. Moreover, there was confusion among students, whether the mathematical model should be fitted to the data or vice versa in the sense that the measurement of the data should be improved in order to fit better to the mathematical model. In this contribution, we present an extension to the aforementioned pilot study to get more insights into students' validation approaches.

Theoretical Background

Mathematical Modelling

Mathematical modelling is usually described as a circular process (e.g. Galbraith et al., 2010). In this contribution, we follow the modelling cycle proposed by Blum and Leiß (2007) describing seven steps in the modelling process (Figure 1). According to Niss (1994), the validation step is the most important step in the modelling-process. In this step students evaluate whether the real results obtained in the modelling process and the used mathematical model itself are adequate with respect to the real situation. As a consequence of this step, the model itself could be revised if it is not found suitable. Unfortunately, validation is very challenging for many students and some students do not see validation as a necessary step (Hankeln, 2020). When students work on modelling tasks, the validation is often absent (Blum & Leiß, 2007) and those students who validate their models often rely on a rather intuitive feeling that their model might be suitable or not (Borromeo Ferri, 2006). In accordance with these results, Borromeo Ferri et al. (2013) report that students scored weakest concerning their validation competence compared to the other modelling sub-competences. Furthermore, within their intervention-study they showed that fostering students' validation-competence is possible even with short interventions like modelling project days.

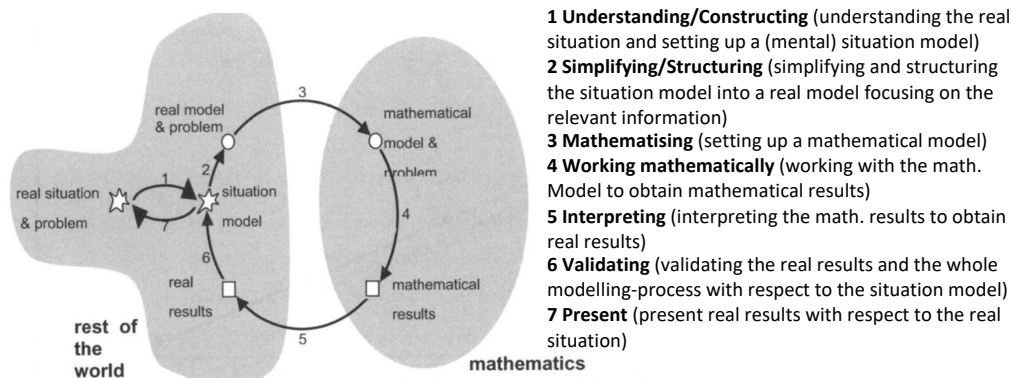


Figure 1: Modelling-Cycle following Blum & Leiß (2007)

Fostering validation with modelling tasks with experiments

In the literature, there exist several concrete teaching ideas for combining experiments and modelling tasks (e.g., Ludwig & Oldenburg, 2007). Halverscheid (2008, p. 226) sees a natural link between experiments and modelling as experiments “represent the *rest of the world* for which mathematical models are built.” Ludwig and Oldenburg (2007) emphasize benefits of experiments for the whole modelling process and especially the validation step, because students can validate their models based on real data instead of theoretical considerations. According to Zell and Beckmann (2009, p. 2216) experiments offer good occasions for validation: “Because of measurement errors the formula is never correct. So it is natural to talk about the correctness and the limitations of the model and its results.”

Even if these considerations propose a high potential of modelling tasks with experiments for validation, empirical research in this field is scarce. So far only a few qualitative studies with small samples have dealt with validation in the context of modelling tasks with experiments. Zell and Beckmann (2009) used modelling tasks in combination with physics experiments and interviewed

secondary students after working on these tasks. Students struggled with accepting the imperfection of their models due to measurement errors at first but eventually were able to validate their models. In contrast, Maull and Berry (2001) observed undergraduate students who worked on a modelling task concerning the cooling off of tea and found that many students were not able to validate their models without help of the lecturer. It seems that using experiments can even hinder the validation process: Carrejo and Marshall (2007) report about an undergraduate course in which students experimented with motion. When talking about the validation of their models, students attributed even systematic shortcomings to measurement errors instead of unfavorable model specifications.

In a pilot study of the MaMEx-project, Geisler (2021) implemented a modelling task concerning the decay of beer-froth in two upper secondary classes. Students performed an experiment and used their gathered data to model the decay by exponential functions. Analyzing students written solutions, Geisler (2021) found that students' acceptance or rejection of a model was very subjective, in the sense that models identified as adequate by the students as well as models that students judged as inadequate showed the same systematic shortcomings (e.g. systematically over- or underestimating the data). These shortcomings have often been justified via measurement errors by the students which is in line with Carrejo and Marshall's (2007) findings. Furthermore, being asked how to improve the fit between data and function, most students stated that they would improve their experimental procedures instead of searching for an alternative model. It seems that students put more trust into their mathematical model than in their data and therefore try to improve data in order to fit the data to the model instead of improving the model with respect to the data. However, the analyzed sample ($N = 19$) in this pilot study was rather small and all students came from the same school.

The current study

Objectives and Research Questions

In this paper, we present an extension to the aforementioned study. We implemented two modelling tasks in four secondary classes to determine whether results are comparable within a larger as well as more diverse sample and with different modelling tasks. Furthermore, we wanted to understand students' ideas behind their approaches to improve their models. Our guiding research question was:

How do students validate their models within modelling tasks with experiments with respect to their gathered data and which approaches to improve models do they propose?

In the following, we give an overview of the used modelling tasks before describing the methods of data collection and analysis.

Design Principles of the Modelling Tasks

All modelling tasks with experiments used in the MaMEx-project follow four design-principles (cf. Galbraith et al., 2010): P1) The contexts of the task should be realistic and use only those physical quantities that are already known by the students. P2) The mathematics should be foregrounded in the whole process. Therefore, experiments should be easy to conduct and take only a few minutes. P3) It should take not too much time to set up a mathematical model but the task should offer relevant occasions for validation. P4) A validation prompt should be implemented since many students do not validate their models spontaneously.

The task “Cold Coffee” (in the following named “CC”) is adapted from Ludwig and Oldenburg (2007). The following served as an introduction to the task:

After brewing, coffee needs some time to cool off in order to be conveniently drinkable. The desired drinking temperature differs from person to person. Model the temperature decrease and evaluate at which time the coffee can be delightfully drunken.

The context is familiar to students. Students are asked to formulate a hypothesis concerning the decrease of temperature before performing an experiment. Students measure the temperature of freshly brewed coffee every minute for 10 minutes. Only well-known quantities (time and temperature) are used in the experiment (P1) and it can be done with everyday material and takes only 10 minutes to perform (P2). After performing the experiment, students were asked to model the cooling using a function. No hint was given which kind of function is suitable. However, setting up a mathematical model is not too complicated (P3) since the temperature can be modelled using an exponential function of the form $f(x) = b \cdot a^x + c$, $b > 0$, $c > 0$, $0 < a < 1$ with x as the time in minutes after brewing the coffee. The parameters can be calculated by using the data from the experiment: c can be estimated as the room temperature and b can be calculated as the difference between the temperature at the beginning of the experiment and the room temperature. One possibility to estimate a is to use two consecutive measured values, e.g. $a = \frac{\text{temp}(1 \text{ min}) - b}{\text{temp}(0 \text{ min}) - b}$. Students could use

GeoGebra to plot graphs of functions and compare them with their data. There are several occasions for validation, since the decrease is not perfectly exponential. Furthermore, students sometimes forget to consider that the coffee cannot cool down under the room temperature, so that their first attempt to model the decrease is often not adequate. As a validation prompt (P4) serves the following subtask:

Compare your function with your measurement-data. Does your function describe the data accurate enough? How could your model be improved?

Besides the CC task, a second modelling task “Stale Beer” (“SB” in the following) was used. In this task students were asked to model the decay of beer-froth and to evaluate the beer quality based on their results (Geisler, 2021). The SB task had the same structure as the CC task, in the sense that the same subtasks (e.g., the same validation prompt) were used. However, the SB task can be considered to be a little easier. The decay of beer-froth can be modelled using an exponential function of the type $f(x) = b \cdot a^x$ and therefore less parameters have to be computed. A detailed description of this task and the related experiment can be found in Geisler (2021).

Methods

Both modelling tasks were implemented in four upper secondary classes (all in year 10) from three different German schools. 71 students (age between 15 and 17 years, 54 % female) voluntarily participated in the study. Based on their last mathematics grades, the students represent a large variety of achievement levels reaching from grade 1 (very good) to grade 6 (not sufficient) (mean: 3.1 – sufficient). All classes had covered exponential functions as well as the calculation of the function’s parameters based on given values some month prior to the implementation of the tasks. Students were not informed that exponential functions could be used to solve the tasks. In every class, half of the students worked on the CC task whereas the other students solved the SB task during a 90 minutes

lesson. Students performed the experiments in pairs of two. Students' answers to the validation prompt served as data in order to answer the research question. Since not all students answered all subtasks, data is available for 64 students. Furthermore, we conducted stimulated recall interviews with six students after working on the task to better understand students' validation ideas and ideas for model improvement. We chose students for the interviews based on their answers to the validation prompt, to achieve a certain variety of students' approaches. Accordingly, we used their answers to the validation prompt as a stimulus in the interviews. Interviews have been recorded and transcribed.

Students' answers have been analyzed using qualitative content analysis (Mayring, 2010). We used an inductive coding guide. All answers were first coded regarding students' judgement whether their model was adequate with respect to their data. Second, students' ideas to improve their models were coded. All answers have been coded independently by two raters and Cohen's kappa has been calculated indicating a very good interrater-reliability of $\kappa = 0.9$.

Results

Since both tasks are very similar in structure, we do not discuss the results divided by tasks. However, numbers concerning the different approaches related to the tasks are provided in brackets. Except of two (SB: 1, CC: 1), all students were able to set up an exponential function as a model for the decay of beer-froth respectively the decrease of the temperature of the coffee. However, most students working on the CC task did not consider the room temperature and set up functions of the form $f(x) = b \cdot a^x$ resulting in models that mostly only describe adequately the temperature during the first minutes of the experiment, as can be seen in the solutions in Figure 2.

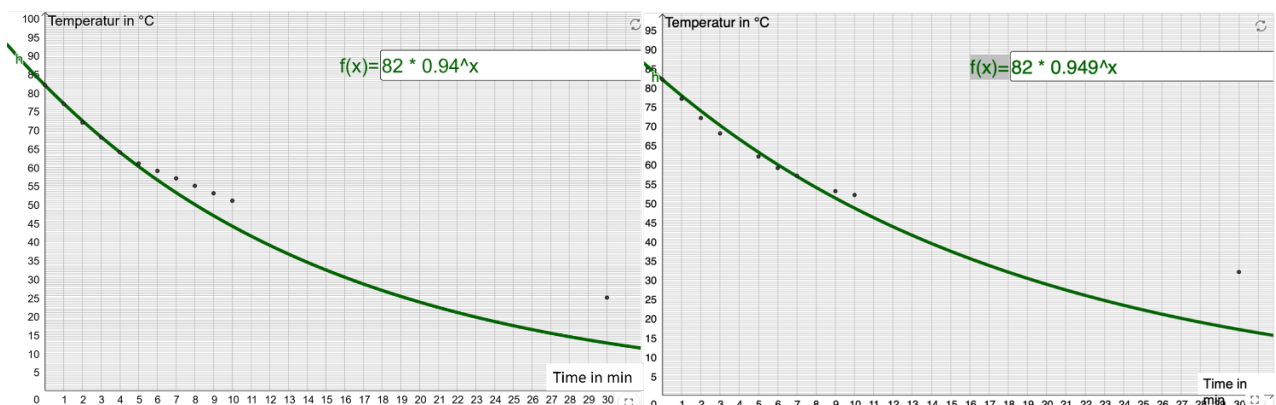


Figure 2: Screenshots of students' GeoGebra files with measured data and related functions – CC task

Model Validation

28 students (SB: 16, CC: 12) stated that they see no or minimal deviations between their model and data, indicating that they consider their model to be adequate (e.g., “*The function describes our measured values quite exact*”, CC). In contrast, 25 students (SB: 13, CC: 12) see substantial deviations between their data and their model and therefore judge their model as not adequate (e.g., “*No, the function is decreasing more steep than our measured values*”, SB). Two of these students wrote that they are satisfied with their model but that they consider their measured values to be “*very imprecise*”. It seems that these students attribute deviations between model and data to their data instead of the specifications of the model. 11 students (SB: 6, CC: 5) gave a more sophisticated

evaluation, stating in which timespans the model is useful to describe their data (e.g., “*The measured values deviate a little bit from the function especially at the beginning. At the end the values match with the function*”, SB). In these answers, first ideas concerning the limitations of models become apparent. These judgements are nearly equally distributed among the two tasks. Furthermore, students’ judgements are very subjective, as can be seen in the solutions in Figure 2. While both models show similar shortcomings (good fit during first minutes but clear deviations later), the left model was considered to be not adequate by students and the right model was considered sufficient.

Approaches for Model Improvement

Five students did not answer the question concerning the improvement of their models. Moreover, ten students (SB: 4, CC: 6) explicitly stated that they have no concrete idea for improving the fit between model and data. The majority of the students ($n = 39$, SB: 22, CC: 17) only gave ideas to improve the experimental procedure instead of working on the model itself. Most of these ideas contain only simple improvements like taking more measurement values or measuring the values more precisely. However, two students argued that one should repeat the experiment several times with partly different materials (“*Take more values for a longer timespan for a more precise model, repeated experiments with different bottles of the same beer*”, SB). Only ten students (SB: 4, CC: 6) wrote ideas for an alternative model for their data, like calculating the parameter “ a ” in another way (e.g., “*It would be helpful to calculate a for all measured values and to take the mean for it*”, CC).

It seems that many students have no concrete idea how an alternative model based on their already measured values can be set up. Furthermore, some students seem to put more trust in their model than in their data. This impression is supported by students’ answers in the stimulated recall-interviews. Out of the six students interviewed, only two (we call them student *A* and student *B*, both worked on the CC task) offered ideas how to improve the function. Student *A* offered the idea to use more than one function (in the sense of a composite function) in order to adequately model different parts of the data: “We recognized that beginning with a certain value it [the coffee temperature] remained constant. [...] Perhaps one can split it into two functions: one for the first part and one for the constant part”. Student *B* followed a different idea that involved using all measured values to set up a function:

- Interviewer: You have thought about how one can improve your function and you wrote “one could calculate a for every single value and take the mean then.” What do you mean by every single value?
- Student *B*: We have values for every timepoint and we could calculate an a for every timepoint and calculate the mean of these a s.
- Interviewer: And what would happen if you do so?
- Student *B*: Then one value would lie exactly on the function and it would be consistent what lies over and under the graph.

Student *B* seems to be aware that not all measured values will lie exactly on a function and that it is desirable that a similar number of values lies over and under the function, that is the function does not systematically over- or underestimate the measured data.

All other students said that they have no idea how to further work on their functions. Two students argued that the model depends heavily on the experimental data and that the experiment should be improved accordingly. These students are aware that measuring new data will result in a slightly different function, as becomes apparent in the dialogue with student *C*, who worked on the SB task:

- Interviewer: Do you have ideas how to improve your model if you already have the measured values? [...]
- Student *C*: So I think the analysis is really dependent on the experiment and so it is more relevant what values we gather from the experiment to have the right analysis.
- Interviewer: Ok, so you say it depends on the data. What would you say would happen to your function if you gather other data, if you measure more often or more precisely?
- Student *C*: I think nothing. It would be the same. It would just be more detailed and one could better see how the function...
- Interviewer: So if you would have measured more precisely it would be the same function but everything would fit better?
- Student *C*: So maybe there would be a different number [for the parameter] but basically it would be the same.

The last two interviewed students, one is student *D* who worked on the SB task, stated that improving the experiments will lead to the same function they already have:

- Interviewer: You wrote that one could use a better measuring cylinder to better measure the change [of the beer-froth]. These are ideas that set in during experimentation. Do you have also ideas how to improve your function if you already have the values and you cannot change them? [...]
- Student *D*: I have no idea!
- Interviewer: And if you would have a better measuring cylinder and you measured the values more precisely, do you think at the end you would have a function that fits better to the measured values?
- Student *D*: I think we would have the same function then, but eventually the measured values would fit better to it.

Discussion

All in all, the results of the study at hand are similar to those of the pilot study of the MaMEx project (Geisler, 2021) and support the findings of Carrejo and Marschall (2007): Most students seem to attribute deviations between their model and the experimental data to measurement errors and primarily suggest to improve the experimental procedure instead of considering alternative models. This is even the case, if systematic shortcomings of the models exist. These results show that integrating experiments in lessons does not guarantee successful validation but experiments could be used as a starting point for fruitful discussions on validation in the classroom. Regarding students' statements in the interviews, two main problems can be identified: 1) Students seem to have not enough knowledge about exponential functions to reconsider their model and 2) there seems to be a lack of understanding concerning the relation of the reality (data) and the model (function) which has to be addressed in mathematics lessons. The latter problem might be due to the perception that the model is solely determined by the data and that gathering good data is the most important step in the whole process. This is remarkable since ones' model depends on both: the data used and the decisions made (e.g., type of function) in order to set up the model. However, in the answers of student *D* a real misconception concerning the relation between data and model comes apparent. Student *D* seems to see his function as the "right" model and rather independent from the actual data. It remains uncertain where this trust in his model comes from. One possibility could be that many students believe that mathematics is always precise and mathematics tasks have only one right answer. These beliefs are quite common among students and have been proven to influence their approaches to mathematics tasks (Schoenfeld, 1992). Measuring students' beliefs concerning the nature of mathematics prior to their work on modelling tasks with experiments might shed light on this hypothesis.

References

- 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 modeling: Education, engineering, and economics* (pp. 222–231). Horwood. <https://doi.org/10.1533/9780857099419.5.221>
- Borromeo Ferri, R. (2006). Theoretical and empirical differentiations of phases in the modelling process. *ZDM*, 38(2), 86–95. <https://doi.org/10.1007/bf02655883>
- Borromeo Ferri, R., Grünewald, S., & Kaiser, G. (2013). Effekte kurzzeitiger Interventionen auf die Entwicklung von Modellierungskompetenzen. In R. Borromeo Ferri, G. Greefrath & G. Kaiser (Eds.), *Mathematisches Modellieren für Schule und Hochschule* (pp. 41–56). Springer. https://doi.org/10.1007/978-3-658-01580-0_2
- Carrejo, D. J., & Marshall, J. (2007). What is mathematical modelling? Exploring prospective teachers' use of experiments to connect mathematics to the study of motion. *MERJ*, 19(1), 45–76. <https://doi.org/10.1007/bf03217449>
- Engel, J. (2010). *Anwendungsorientierte Mathematik: Von Daten zur Funktion*. Springer. <https://doi.org/10.1007/978-3-540-89087-4>
- Galbraith, P. L., Stillman, G. A., & Brown, J. (2010). Turning ideas into modeling problems. In R. Lesh, P. L. Galbraith, C. Haines & A. Hurford (Eds.), *Modeling students' mathematical modeling competencies* (pp. 133–144). Springer. https://doi.org/10.1007/978-1-4419-0561-1_11
- Geisler, S. (2021). Data-baseball modelling with experiments - Students' experiences with model-validation. In M. Inprasitha, N. Changsri & N. Boonsena (Eds.) *Proc. 44th Conf. of the Int. Group for the Psychology of Mathematics Education, Vol. 2* (p. 330–339). PME.
- Hankeln, C. (2020). Mathematical modeling in Germany and France: A comparison of students' modeling processes. *Educational Studies in Mathematics*, 103, 209–229. <https://doi.org/hnw7>
- Halverscheid, S. (2008). Building a local conceptual framework for epistemic actions in a modelling environment with experiments. *ZDM*, 40(2), 225–234. <https://doi.org/dj2v4b>
- Ludwig, M. & Oldenburg, R. (2007). Lernen durch Experimentieren. Handlungsorientierte Zugänge zur Mathematik. *mathematik lehren* 141, 4–11.
- Mauß, W., & Berry, J. (2001). An investigation of student working styles in a mathematical modelling activity. *Teaching Mathematics and Its Applications*, 20(2), 78–88. <https://doi.org/10.1093/teamat/20.2.78>
- Mayring, P. (2010). *Qualitative Inhaltsanalyse*. Beltz. <https://doi.org/c7ksqg>
- National Governors Association Center for Best Practices and Council of Chief State School Officers. (2010). *Common core state standards for mathematics*. CCSSO.
- Niss, M. (1994). Mathematics in society. In R. Biehler, R. W. Scholz, R. Sträßer & B. Winkelmann (Eds.). *Didactics of mathematics as a scientific discipline* (pp. 367–378). Kluwer.
- Schoenfeld, A. H. (1992). Learning to think mathematically. In D. Grouws (Ed.), *Handbook for Research on Mathematics Teaching and Learning* (pp. 334–370). Macmillan.
- 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.