# PROCEEDINGS OF THE 47th CONFERENCE OF THE INTERNATIONAL GROUP FOR THE PSYCHOLOGY OF MATHEMATICS EDUCATION

Auckland
Aotearoa New Zealand
July 17-21
2024

## EDITORS

Tanya Evans
Ofer Marmur
Jodie Hunter
Generosa Leach
Jyoti Jhagroo

## VOLUME 3

Research Reports
(H – O)

# PROCEEDINGS OF THE 47th CONFERENCE OF THE INTERNATIONAL GROUP FOR THE PSYCHOLOGY OF MATHEMATICS EDUCATION

Auckland
New Zealand
July 17-21
2024

## EDITORS

Tanya Evans
Ofer Marmur
Jodie Hunter
Generosa Leach
Jyoti Jhagroo

## VOLUME 3

Research Reports
(H – O)

PME-47

RETHINKING MATHEMATICS EDUCATION TOGETHER

**Cite as:**

Evans, T., Marmur, O., Hunter, J., Leach, G., & Jhagroo, J. (Eds.) (2024). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3). Auckland, New Zealand: PME.

Website: https://events.massey.ac.nz/pme-47-conference/

Proceedings are also available on the IGPME website:
http://www.igpme.org



ISBN: 978-1-0670278-3
ISSN: 0771-100X

Logo designed by Jason Lamontanaro

**TABLE OF CONTENTS – VOLUME 3**

# PRESERVICE TEACHERS' USE OF TEACHER MOVES THAT PROVIDE SENSE-MAKING OPPORTUNITIES TO STUDENTS WHEN THEY IMPLEMENT NUMBER TALKS

Simon Byeonguk Han[1] and Byungeun Pak[2]

[1]Portland State University, [2]Utah Tech University

*Supporting students to make sense of mathematical ideas is crucial in mathematics classrooms. Number Talks center students' mathematical ideas, which is suitable for pursuing students' sense-making of mathematics. However, there are a very limited number of empirical studies regarding the efficacy of Number Talks. Also, teacher moves during the Number Talks have not been studied yet. In this study, we analyzed videos of 22 Preservice Teachers' 48 Number Talks from grades 3 to 5. We examined what teacher moves were used and their frequencies in the Introducing and Idea Sharing phase. We identified nine teacher moves, either constructive or interactive, and their frequencies, which could potentially support students' sense-making, within three teaching practices in NTs.*

## INTRO AND RATIONALE

In mathematics classrooms that center students' sense-making, students "have experiences that enable them to connect new learning with prior knowledge and informal reasoning" (NCTM, 2014, p. 9). Sense-making happens when students engage in certain (meta)cognitive activities, which are reflected in overt behaviors (Fiorella, 2023). Chi and Wylie (2014) proposed four modes of cognitive engagement and detailed students' overt behaviors reflecting each mode during learning: passive, active, constructive, and interactive. Number Talks (NTs) are 5-15 instructional routines that center students' mathematical ideas. Despite the popular use of NTs in math classrooms in the U.S., there are a very limited number of empirical studies on the efficacy of NTs. Similarly, there was no research on the efficacy of NTs presented at PME at least for the last ten years, We pay attention to overt behaviors in NTs to highlight what Preservice Teachers (PSTs) may do to promote students' sense-making. Our research question is: *What types of teacher moves PSTs use in NTs to support students' overt behaviors in constructive and interactive mode?*

## LITERATURE REVIEW AND FRAMING

Chi and Wylie (2014) argued that in a *passive mode*, the lowest level of engagement, learners are "oriented toward" and receive "information from the instructional materials without overtly doing anything else related to learning" (p. 221). In an *active mode*, learners engage with the instructional materials with "some form of overt motoric action or physical manipulation" (p. 221). In a *constructive mode*, learners "generate or produce additional externalized outputs or products beyond what was provided in the learning materials" (p. 222). In an *interactive mode*, the highest engagement mode, learners engage in interpersonal activities. Fiorella (2023) argued



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 1–8). PME.

that learning activities involving a constructive or interactive mode support students' sense-making compared to an active or passive mode.

NT is a 5 to 15-minute instructional routine in mathematics classrooms. It consists of four phases, (1) Introducing, (2) Collecting Answers, (3) Idea Sharing, and (4) Closing (Humphreys & Parker, 2015; Parrish, 2014). The benefits of participating in NTs include students clarifying and explaining their thinking, exploring mathematical relationships, understanding many possible ways to solve a problem, and learning to trust their reasoning (Han & Thanheiser, 2021; Sun et al., 2018; Woods, 2022). In the Introducing phase, students privately and quietly think about their answers and strategies (private thinking time). The Idea Sharing phase begins with the teacher creating a public record of the strategies shared by students. Rumsey et al. (2019) and Thanheiser and Melhuish (2023) referred 'public records' of students' mathematical thinking as written notes of students' mathematical ideas shared by students. Creating and working with public records is recognized as one of the teaching practices associated with instructions that value students' sense-making (Thanheiser & Melhuish, 2023).

Studies on overt behaviors reflecting sense-making (Chi & Wylie, 2014; Fiorella, 2023) did not use NTs. However, in the Introducing and Idea Sharing phase, students are encouraged to construct their strategies when they are asked to think alone about strategies and share their ideas with peers and teachers (Humphreys & Parker, 2015; Pak et al., 2023; Parker & Humphreys, 2018; Parrish, 2014), which aligns with the constructive and interactive mode in Chi and Wylie's (2014) categorization. Pak et al. (2023) demonstrated that novice teachers may not go beyond serial sharing in their NTs. The teachers interacted with a single student in a one-on-one, which may leave many other students in a potentially passive mode of engagement. To make NTs a potentially effective way for students to engage in sense-making, we propose supporting students to engage in a constructive and/or interactive mode during NTs in the Introducing and Idea Sharing phase is important.

Studies on mathematical discourse identified specific teacher moves. In such classrooms, teachers ask questions to elicit and clarify student thinking (Herbel-Eisenmann et al., 2013), justify their reasoning (Thanheiser et al., 2021), give enough wait time (Chapin et al., 2013), and revoice what students say to confirm the accuracy of teachers' understanding of mathematical ideas (Kazemi & Hintz, 2014), or press students to detect mathematical errors (Pak et al., 2023). These teacher moves support students to have sense-making opportunities since it invite students to be in a constructive and interactive mode of cognitive engagement.

In the study working with 12 elementary school teachers, Franke et al. (2015) found six *teacher invitation moves* that included asking students to (1) explain someone else's solution, (2) discuss differences between solutions, (3) make a suggestion to another student about their idea, (4) connect their ideas to other students' ideas, (5) create a solution together with other students, and (6) use a solution that was shared by another student publicly. Pak et al. (2023) illustrated six teacher moves that might support students to engage in and contribute to each other's ideas: Inviting (directly inviting students to question the strategy sharer), prompting interpretation (asking another student to offer their reasoning regarding how the initial strategy works), asking for (dis)agreement (asking students to (dis)agree a student's ideas with follow-up questions), cueing an error (detecting error and challenging the strategy sharer), guiding (generating the reasoning to draw attention to something specific in the strategy), and repeating (asking students to revoice what another student says).

## METHODS

This study is part of the project that aims to investigate how PSTs learn to implement ambitious and equitable teaching practices (NCTM, 2014). It was conducted by the first author at an elementary teacher preparation program (four semesters long) in a Southwest State in the U.S. in Spring 2023. The subjects of this study were in their second semester, a cohort of 29 PSTs, who took a mathematics methods course for elementary school teachers. 25 PSTs agreed to participate in the study. Each of them was assigned to a classroom between the third and fifth grades. None of them were exposed to NTs before. In the course, PSTs had multiple opportunities to learn about ways to implement NTs via instructional activities in class and the NT project. PSTs watched videos of experienced teachers' NTs and discussed, read short readings about NTs, and engaged in an instructional activity in which they came up with multiple ways to compare multiple strategies working with their peers in groups. As a whole, the instructional activities were intended for the PSTs to apply what they learned to the NT project. Then, the PSTs were required to complete the NT project (see Figure 1). Each PST electronically submitted their NT plan, revised plan, a video of their enacted NTs, and reflection/analysis paper to a learning management system.

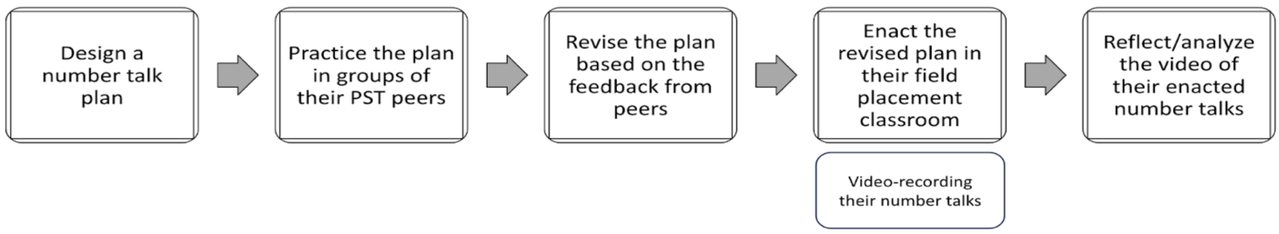



Figure 1: Phases of the Number Talk Project

We excluded 3 videos and analyzed 22 videos of NT implementations from our data sources due to the low audio and video quality. The number of NTs in each video ranged from one to four. Our analysis took three steps. First, each author watched all the videos independently and identified three teaching practices in the Introducing and Idea Sharing: (1) offering a private thinking time; (2) creating public records of an individual student's math thinking; and (3) engaging multiple students in peers' strategies. Second, we identified teacher moves associated with each teaching practice. To code teacher moves related to each teaching practice, we began with teacher moves identified in the related studies (Pak et al., 2023; Franke et al., 2015). Each researcher individually coded the video and remained open to any emerging codes. We tried to identify all teacher moves and discuss each teacher move in terms of whether and how each teacher move promotes students' sense-making opportunities and whether each teacher move helped students engage in a constructive or interactive mode. After discussions, we came up with the nine teacher moves that PSTs used to promote students' sense-making opportunities within the three teaching practices (see Table 1). Third, for the nine teacher moves, we examined their frequencies by PSTs, NTs, and strategy episodes and identified some exemplar excerpts for each teacher move. The frequency was counted by talk turns but excluded one-word responses.

## FINDINGS

We identified 152 strategy episodes (i.e. students shared 152 strategies) in 48 NTs by 22 PSTs. We found that all NTs were done in a serial sharing manner which involved multiple teacher moves. We captured two teacher moves during the private thinking time, five teacher moves related to public records, and two teacher moves related to engaging multiple students. Table 1 shows all teacher moves and their frequencies.

| Teacher Practices | Teacher Moves (Total Incidents of Each Teacher Move) | Strategy Episode (152 episodes) | NTs (48 NTs) | PSTs (22 PSTs) |
|---|---|---|---|---|
| Offering a Private Thinking Time | Encouraging Students to Do Mental Math (23) | NA | 20 NTs | 13 PSTs |
| | Encouraging Students to Use Hand Signals (31) | NA | 28 NTs | 17 PSTs |
| Creating Public Records of an Individual Student's Math Thinking | Revoicing What a Student Said (147) | 95 Episodes | 44 NTs | 22 PSTs |
| | Asking Students Questions for Clarification (54) | 43 Episodes | 27 NTs | 17 PSTs |
| | Asking Students to Justify Their Reasoning (14) | 7 Episodes | 5 NTs | 3 PSTs |
| | Responding to Students' Self-correction on Mathematical Errors (4) | 4 Episodes | 3 NTs | 2 PSTs |
| | Mentioning a Brief Connection between Strategies (3) | 3 Episodes | 3 NTs | 3 PSTs |

| | | | | |
|---|---|---|---|---|
| Engaging Multiple Students in Peers' Strategies | Recapping a Student's Strategy for the Whole Class (10) | 10 Episodes | 9 NTs | 8 PSTs |
| | Asking Students to Make Connections Between Mathematical Ideas (1) | 10 Episodes | 1 NT | 1 PST |

Table 1: Nine Teacher Moves and Their Frequencies

**Finding 1: Teacher Moves during Private Thinking Time**

We found two teacher moves in 48 private thinking times. The first teacher move was *encouraging students to do mental math* (13 PSTs used it 23 times in 20 NTs). For example, PST18 encouraged students to show her their thinking by saying, *"Show me you are thinking. Everybody should be thinking ... before you have your answer... one or the other"* in her second NT. The second teacher move was *encouraging students to use hand signals* (17 PSTs used it 31 times in 28 NTs). For example, PST7 mentioned, *"Where should you put your hand if you have an answer?"* in her first NT.

**Finding 2: Teacher Moves Regarding Creating Public Records of An Individual Student's Mathematical Thinking**

We grouped five teacher moves that the PSTs used when they created public records into two groups based on the frequencies. More frequently used teacher moves were (1) *revoicing what a student said* (22 PSTs used it 147 times in 95 strategy episodes in 44 NTs) and (2) *asking students questions for clarification* (17 PSTs used it 54 times in 43 strategy episodes in 27 NTs). Less frequently used teacher moves were (1) *asking students to justify their reasoning said* (3 PSTs used it 14 times in 7 strategy episodes in 5 NTs), (2) *responding to students' self-correction on mathematical errors said* (2 PSTs used it 4 times in 4 strategy episodes in 3 NTs), and (3) *mentioning a brief connection between strategies said* (3 PSTs used it 3 times in 3 strategy episodes in 3 NTs). Table 2 shows examples (highlighted in bold) of each teacher move.

| Teacher Moves | Example |
|---|---|
| Revoicing What a Student Said | Student: I know that 15 plus 15 equals 30. And because 3 times 3 equals 9, 3 times 30 equals 90.<br>PST9: So, **3 times 3 is 9**... You knew that |
| Asking Students Questions for Clarification | Student: 3 times 30 equals 90.<br>PST9: 3 times 30 equals 90. Okay. So then, **how did you get 6**? |
| Asking Students to Justify Their Reasoning | Student: I divided it by 10.<br>PST4: Divided what by 10?<br>Student: 70.<br>PST4: Okay. **Why did you divide 70 by 10?** |

| | |
|---|---|
| | Student: To make multiplication easier. |
| Responding to Students' Self-correction on Mathematical Errors | Student: And then I added... Wait…<br>PST7: Then you added?<br>Student: Yeah. No. take off…<br>PST7: You subtracted 3 from that<br>Student: … [Still thinking].<br>PST7: That is alright.<br>Student: Because I added 3 at the start. You have to take 3 off.<br>PST7: Hmm. Yes, but you added 3 to 67 and made it 70. If you didn't take away 3 from 232, we still have an extra 3. **I see where you are confused.** |
| Mentioning a Brief Connection Between Strategies | PST3: Okay, so you took this one and **[pointing out another strategy recorded on the board] you kind of did it this way**, right? Okay, awesome. |

Table 2: Examples of Each Teacher Move in Creating Public Records

**Finding 3: Engaging Multiple Students in Peer's Strategies**

We found two teacher moves when PSTs tried to engage multiple students in each other's strategies: (1) *recapping a student's strategy for the whole class* (8 PSTs used it 10 times in 10 strategy episodes in 9 NTs) and (2) *asking students to make connections between mathematical ideas* (1 PST used it once in 1 strategy episode in 1 NT). This teaching practice was built on public records of students' strategies. There were two types of recapping, either the teacher recaps the student's strategy from beginning to end or labels the student's strategy (e.g. doubling and halving). We captured one episode of *asking students to make connections between mathematical ideas*. In PST15's second NT, the PST asked the whole class why they think the two different problems (the first and second NT) resulted in the same answer ("Can anyone tell me why you think 4 times 64 would be the same answer as 8 times 32?").

**DISCUSSIONS**

In summary, we found 9 teacher moves used by the PSTs in NTs that possibly support students' sense-making as they engaged in three teaching practices. In this section, we provide discussion points along with corresponding findings. First, PSTs used teacher moves during private thinking time, which is not aligned with what experts on NTs recommend teachers do. Experts suggest that students should not be disturbed as they think about answers and strategies (e.g., Humphreys & Parker, 2015; Parrish, 2014). It is unclear to us whether those teacher moves disrupted students' thinking or helped them to focus but it was not in the scope of this study. For future study, asking students how they felt about those teacher moves could help us to understand it.

Second, the finding regarding public records of individual students' thinking shows a strong potential to help students engage in sense-making. In NTs, PSTs not merely created public records but used multiple teacher moves that possibly supported students' idea sharing. However, although the PSTs were exposed to a variety of teacher moves to promote students' sense-making throughout the course before they enacted NTs and were encouraged to incorporate such teacher moves in their NTs, they utilized only certain types of teacher moves. However, since we only observed one NT implementation from each PST and we do not know how experts on NTs use teacher moves in their NT, we cannot demonstrate whether it was due to PST's lack of experience in teaching and/or NTs. Therefore, we need to investigate both how to support PSTs to integrate a variety of teacher moves into NTs and how NT experts utilize teacher moves in their NTs in the future.

Third, we were able to see only a limited number of teacher moves concerning multiple student engagements happening. Facilitating mathematical discussions in ways to engage students deeply with each other's strategies is known as challenging even for experienced reform-oriented teachers (Kazemi & Stipek, 2001). A previous study showed that NTs enacted by beginning teachers in their early career of teaching rarely included such ways to facilitate mathematical discussions (Pak et al., 2023). Therefore, our finding that shows only a few instances of teacher moves for multiple students' engagement is not surprising. However, for similar reasons in the second point, we do not have evidence of whether it was because of a lack of experience. Future research needs to investigate how novice teachers' use of teacher moves for multiple student engagements is changed over time and how experienced teachers utilize it in their NTs.

Overall, since we only studied PSTs' teaching moves and did not inquire about students, we do not know whether sense-making happened to students through those 9 teacher moves. We will study both teachers and students in the same classroom and explore the relationship between sense-making opportunities provided by the teacher and students' sense-making in NTs in the future.

## References

Chapin, S. H., O'Connor, C., & Anderson, N. C. (2013). *Classroom discussions: Using math talk to help students learn* (2nd ed.). Math Solutions

Chi, M.T.H., & Wylie, R. (2014). The ICAP framework: linking cognitive engagement to active learning outcome. Educational Psychologist, 49(4), 219-243.

Fiorella, L. (2023). Making sense of generative learning. *Educational Psychology Review, 35*, 1-42.

Franke, M. L., Turrou, A. C., Webb, N. M., Ing, M., Wong, J., Shin, N., & Fernandez, C. (2015). Student engagement with others' mathematical ideas: The role of teacher invitation and support moves. *The Elementary School Journal*, *116*(1), 126–148.

Han, S. B., & Thanheiser, E. (2021). Number talks in asynchronous online classrooms for more equitable participation and as formative assessment of student thinking. In K. Hollebrands, R. Anderson & K. Oliver (Eds.), *Online learning in mathematics education* (pp. 147-166), Springer.

Herbel-Eisenmann, B. A., Steele, M. D., & Cirillo, M. (2013). (Developing) teacher discourse moves: A framework for professional development. *Mathematics Teacher Educator*, *1*(2), 181–196.

Humphreys, C., & Parker, R. (2015). *Making number talks matter: Developing mathematical practices and deepening understanding, grades 4-10*. Stenhouse Publishers.

Kazemi, E., & Hintz, A. (2014). *Intentional talk: How to structure and lead productive mathematical discussions*. Stenhouse Publishers.

Kazemi, E., & Stipek, D. (2001). Promoting conceptual thinking in four upper-elementary mathematics classrooms. *The Elementary School Journal*, *102*(1), 59–80. National Council of Teachers of Mathematics. (2014). *Principles to actions: Ensuring mathematical success for all*.

Pak, B., Cavana, J. M., & Jackson, B. E. (2023). The relationship between number talks and ambitious instruction: Learning from beginning teachers. *Mathematical Thinking and Learning*, 1-24.

Parker, R. E., & Humphreys, C. (2018). *Digging deeper: Making number talks matter even more, grades 3-10*. Stenhouse Publishers.

Parrish, S. D. (2011). Number talks build numerical reasoning. *Teaching Children Mathematics*, *18*(3), 198–206.

Parrish, S. D. (2014). *Number talks: Helping Children build mental math and computation strategies, grades K-5, updated with common core connections*. Math Solutions.

Rumsey, C., Guarino, J., Gildea, R., Cho, C.Y., & Lockhart, B. (2019). Tools to support K-2 students in mathematical argumentation. *Teaching Children Mathematics, 25*(4), 208-217.

Sun, K. L., Baldinger, E. E., & Humphreys, C. (2018). Number talks: Gateway to sense making. *The Mathematics Teacher*, *112*(1), 48–54.

Thanheiser, E., Melhuish, K., Sugimoto, A., Rosencrans, B., & Heaton, R. (2021). Networking frameworks: a method for analyzing the complexities of classroom cultures focusing on justifying. *Educational Studies in Mathematics, 107*, 285-314.

Thanheiser, E., & Melhuish, K. (2023). Teaching routines and student-centered mathematics instruction: the essential role of conferring to understand student thinking and reasoning. *The Journal of Mathematical Behavior, (70),* 1-19.

Wood, T., Williams, G., & McNeal, B. (2006). Children's mathematical thinking in different classroom cultures. *Journal for Research in Mathematics Education*, *37*(3), 222–255.

Woods, D. M. (2022). Building a math-talk learning community through number talks. *Journal of Mathematical Behavior, 67*, 1-21.

# A FRAMEWORK FOR ANALYZING LONG-TERM EARLY ALGEBRA PROGRESSION IN TEXTBOOK SERIES

Ola Helenius and Linda Marie Ahl

University of Gothenburg, Sweden

*We present a two-dimensional framework capable of characterizing the algebra content in textbook series spanning over at least all years of comprehensive schooling. Our framework extends well-known previous work and subdivides school algebra into algebra classes such as structure, operations on non-numerical symbols, functional thinking, and patterns. In each class, we characterize the presented content according to the explicitness levels potential, formal and explicit. We examine two book series, both spanning nine years of schooling, and in one series, two versions for grades 1-3. Results include a radically different focus on algebra in the middle grades in the two series and overarching trends that algebra content is well spread out over school years early but tends to come in bursts in later school years.*

## INTRODUCTION

Algebra is one of the significant branches of mathematics and an essential part of most school mathematics curricula. Historically, school algebra was in many countries introduced when children were about 12 years of age, but somewhere around 1980, a movement that many now call the Early Algebra Movement began to rise, clustered around the idea of introducing some form of algebra in school mathematics for children 6-12 years of age. (Kieran et al., 2016). While many research reports exist describing variants of early algebra, it needs to be clarified to what extent early algebra permeates typical school mathematics, for example, represented in textbooks. The research question we explore is: *How can the strands of early algebra and algebra in textbooks for compulsory school be characterized?*

To examine our research question, we will need an analytical framework capable of characterizing the algebra content in textbook series spanning over at least all years of comprehensive schooling. Such a framework will be described in the next section. Research on the teaching and learning of algebra, and early algebra in particular, is a vast field. Still, fortunately, it is also a field where many summarizing efforts have been made. Our theoretical framework is based on a review of previous work focusing on earlier conference reports, summaries, and reviews (e.g., Kieran, 2018; Kieran et al., 2016). This effort has brought us to distinguish four classes of school of algebra and early algebra that we characterize below.

## THEORETICAL AND ANALYTICAL FRAMEWORK

Of the many ways school algebra and early algebra have been subdivided, we will position our contribution by relating it to the discussion by Kieran (2018), who discusses early algebra in terms of the two categories: algebra as structure in numbers



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 9-16). PME.

and numerical operations and algebra as structure in figural patterns and functions. Blanton and colleagues (2018), building on work by Kaput (2008), use the three categories of generalized arithmetic, equivalence, expressions, equations, inequalities (EEEI), and functional thinking. In relation to Kieran's description, we think it makes sense to look at functional thinking and figural patterns as two separate classes. Functional thinking, comprising the use of variables, covariation, and some aspects of modeling, has a distinguished history connected to school algebra. The role of (the structural aspect of) patterns that are not only number patterns has been highlighted in several large research projects by Mulligan and colleagues (Mulligan et al., 2006).

Blanton and colleagues (2018) decompose the structure in numbers and numerical operations highlighted by Kieran into all things related to equality (EEEI) and the category generalized arithmetic accounting for all other arithmetic structures. While equality and similar relations are undoubtedly important, we prefer to make a different subdivision. We like to distinguish when arithmetical knowledge is generalized or transferred to include working with symbols that are not numerical, like when dealing with equations or expressions. However, we also want to explicitly distinguish when algebraic techniques (that may or may not include symbols) are used to shed light on the structure of (mainly) arithmetic objects, operations, and relations, which gives us four types of algebraic content that are summarized below.

**Algebra as structure**

With *algebra as structure*, we mean ways of discovering, explicating, and symbolizing general relationships among, mainly, arithmetic operations (Freudenthal, 1983; Kieran, 2018; Mason, 2018). 8=6+2 and therefore 8-2=6 and 8-6=2. There is nothing special about 8, 6, and 2 here, so generally, a+b=c, c-a=b, and c-b=a all hold or are all false. Such relations between addition and subtraction can be exploited in classes of experientable situations. Figure 1 shows how schematic iconic imagery can also be used to exploit the mentioned relationship of operations at different levels of generality. Relationships between multiplication and division and between addition and multiplication (i.e., the distributive property), as well as properties of equality, can likewise be explored both by employing experientable classes of situations, iconic schematic imagery, or formally as relations in symbol systems such as those axioms in modern algebra that define a commutative ring with multiplicative identity, or a field.

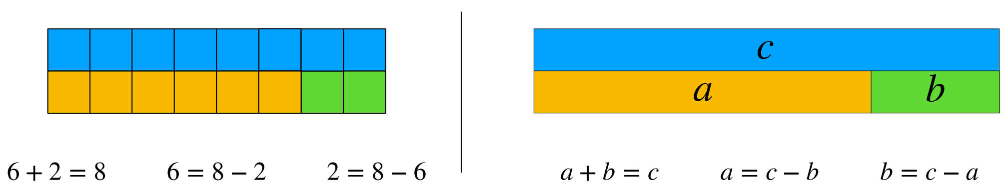



Figure 1. Schematic and symbolic explications of a relationship between addition and subtraction at various levels of generality

**Algebra as operations on non-numerical symbols**

Arithmetic operations on symbols that might stand for a numerical entity but that are not themselves numerical symbols in some sense form the original aspect of school algebra. Typically, a symbol that may be a letter or another token is introduced in various ways and operated on as if it were a number (e.g., Radford, 2022), like when solving linear equations. We also consider that dealing with school tasks like 8+_=3 belongs in this category, as well as tasks like simplifying 4a+4+2a+3.

**Algebra as variables and functional thinking**

Most mathematicians (like the first author of the present paper) would probably argue that the study of functions is not a subfield of algebra, even though overlap exists. Yet, there have been strong arguments from the 1980s and forward for including functional thinking as an aspect of school algebra, which is also represented in national or regional curricula from different parts of the world (Kieran, 2004). Functional thinking is centered around using a symbol that can stand for any number (for which the function is defined), i.e., act as a variable. A central aspect of functional thinking understanding covariation between variables and how it can be graphically represented. In functional thinking, we also include modeling, i.e., when a function is created to represent some (typically) non-mathematical phenomenon.

**Algebra as patterns**

Working with figural or geometric patterns has been highlighted as a way to get acquainted with certain forms of algebraic thinking. Tasks in this class typically involve continuing and representing growing or repeating patterns (Kieran, 2018; Mulligan et al., 2006). Sometimes, tasks involve expressing the number of elements in different instances of a growing pattern. When such tasks involve finding a function for the number of elements in the *n*:th figure, we will classify it as also belonging to the *functional thinking* class. Figural images explicitly used to represent equations are not included in this class but in the class *operations on non-numerical symbols*.

**A dimension of level of explicitness**

Because we wanted to be able to follow the character and distribution of different classes of algebra content over the full spectrum of comprehensive schooling (nine years in Sweden), in addition to the four classes of algebra we presented above, we also devised a classification according to three levels of explicitness in which the algebra content can be presented, experienced and explicated; *potential*, *formal* and *explicit*. We take *potential* to mean when some task or representation can form the basis for some type of algebraic reasoning but where there are no formal indications of algebra. A typical example is 3+_=5, which can be thought of arithmetically as what should be added to 3 to get 5, but where _ also can indicate an unknown number, which is a typical algebraic entity. We take *formal* to mean when the algebraic relationships in play are put in focus by using letters or other non-numerical symbols or by other means. We would, for example, consider exploring the relation between addition and subtraction through iconic schematic imagery, together with corresponding arithmetical expressions like the left side of Figure 1, as formal, even in the absence of non-numerical symbols. We take *explicit* to mean when the algebraic relationships as such are made the focus of attention, like when the distributive property in algebraic form is examined by means of rectangular shapes standing for multiplication, or like in the right side of Figure 1.

## METHOD

We obtained two series of Swedish mathematics textbooks for grades 1-9. For series A, grades 1-3, we examined one version published around 2010 and one published around 2020, which we designate $A_1$ and $A_2$. Because each of the books analyzed formally is a separate reference and that referencing all 36 analyzed books would generate several pages of references, we only give the following indication. $A_1$ is *Matte Direkt Safari, and A 4-6 is Matte Direkt Borgen,* both initially published by Bonniers and later acquired by the publisher Sanoma Utbildning. $A_2$ is the series *Matte Direkt Triumf* published by Sanoma Utbildning. The first author is the same in these three series. Series A 7-9 is *Matte Direkt* published by Sanoma Utbildning. Series B is *Favorit Matematik* in the version *Mera,* second edition for grades 1-6, and *Favorit Matematik* for grades 7-9, all published by Studentlitteratur. All series consist of two books for each grade 1-6 and one book each for each grade 7-9. With our framework in mind, we first analyzed all books page by page, looking for any sign of algebra in either instructions, fact-texts, or tasks. In total, 4133 pages were analyzed. Any page with algebra content was noted and scanned. In a second round, we reviewed the scanned pages. We classified the algebra content into one (or possibly several) categories in our theoretical framework, including if we saw the content as *potentially* algebraic, *explicitly* algebraic, or *formally* algebraic. Because we were interested in how strands of algebra developed over time, we devised a timeline representation based on the approximation that page $n$ in a book with $N$ pages covering one term (or year) was dealt with at a relative time point $n/N$ in that term (or year). Because we were not interested in a detailed sequencing of content, we think this approximation is reasonable. We did not categorize content related to the 10-base positions system and how it can be used for calculations as algebra content unless some other algebraic aspects were dealt with. Even though programming is formally included in the algebra section of the Swedish national curriculum document, we did not categorize programming content. Adjacent pages with the same content were only counted as one algebra unit.

## RESULTS

Our results are presented in Figure 2 as timelines for the analyzed book series, with symbols for the different classes of algebra content. While the graphical representation in Figure 2 is quite dense and complex, it illustrates several important points in how algebra is distributed over school years. For each series, the top row with circles, closest to the timeline, indicates *algebra as structure.* The second row with squares indicates *algebra as operations on non-numerical symbols*. The third row with triangles indicates *functional thinking*. The last row with stars indicates algebra as a structure in *geometric patterns*. For each type, the shade of the symbol indicates the level of explicitness, with white indicating the class *potential*, grey the class *formal,* and black the class *explicit*. We remind the reader that our main aim is to examine if our framework works in the sense of being able to elicit differences in how algebra is treated between textbook series and over time, both within and between textbook series. Of a plethora of possible phenomena visible in the representation in Figure 2, we have picked out six results to present and later discuss both in the light of the framework's functioning and in the light of school algebra.

1. There is an unsurprising general shift from *potential* in earlier grades to *formal* in higher grades, which happens earlier in textbook series A than B, indicated by more grey and less white symbols toward higher grades. There are also very few instances of *explicit* algebra (black).

2. In the lower grades, the algebra content is relatively evenly spread out over the school years, most emphasized in grade 1, and then gradually becomes more and more compartmentalized into specific sections in later grades. This shift happens later in series B. In series $A_1$, the compartmentalization trend can already be seen in grade 2. The trend in series B is very strong in grades 8 and 9 with temporally local but dense clusters. Grade 9 in series A also has such local clusters but much less algebra content overall.

3. The newer series $A_2$ has far more and more evenly spread algebra content and also some content in the classes' *functional thinking* and *patterns*, which is completely missing from the older series $A_1$. $A_2$ also has a more evenly spread out algebra presence than B, while the coverage of the four classes is similar between $A_2$ and B.

4. The A series has a very sparse algebra coverage in grades 4 and 5, and in practice, the books come out as almost algebra-free, in stunning contrast to $A_2$ in earlier grades and to B across grades 4-6.

5. The class *structure* is strong in grades 1-3 but is then thinned out, particularly in series B.

6. In general, *algebra as an operation on non-numerical symbols* has the strongest presence, closely followed by *structure* (see point 5). *Functional thinking* comes third, while there are quite a few instances of the class (geometric) *patterns* overall.

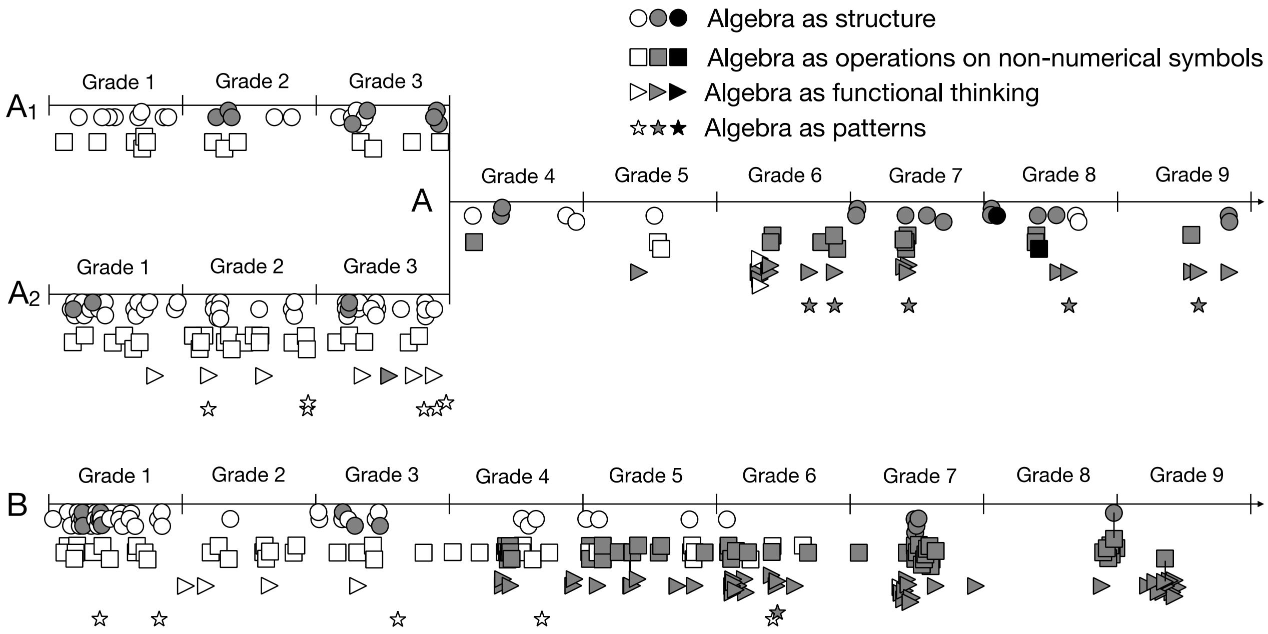



Figure 2. Textbook series A and B algebra content by class and level of explicitness. White symbols represent potential, grey formal and black explicit algebra content

## DISCUSSION

A central point elicited by our results is that the early algebra movement (Kieran et al., 2016) has impacted the textbooks we analyzed, as indicated by the explicit and evenly spread-out algebra content in the early grades (point 2 in the results). The strong early algebra presence can result from our framework being quite generous in identifying algebra in its potential form. The more substantial algebra presence in the newer $A_2$-series books (point 3) may indicate that the early algebra trend has put a change pressure on the textbook market. Assuming such a change pressure, the low algebra content in series A in grades 4-6 might result from these books being initially produced at an earlier point in time. An increased algebra presence in grades 1-3 could be partly due to more emphasis on algebra in the early grades in the 2011 and 2022 national curriculum document revisions. The $A_2$ series came out before 2022, but the national curriculum revision work is relatively open, and publishers generally keep themselves well-informed to have books ready when the new curriculum comes into effect.

The way the algebra content gets compartmentalized in later grades (point 2) can, however, indicate that potential benefits of including algebra early are not taken advantage of. Algebra, equations and functions are dealt with as specific content instead of like overarching mathematical techniques. A reason to focus on early algebra is not only to prepare for formal algebra but also to develop algebraic thinking, to help think of mathematical matters in general in a more structured way (Kaput 2008). We found that the potential ground laid for algebraic thinking in grades 1-3 is not followed up. In series A, algebra almost disappears in grades 4-5 (point 4); in series B, the structural aspect almost disappears from grade 4 and onwards (point 5). The algebra presence shifts to a more classic focus on *algebra as operations on non-numerical symbols* (expressions and equations) and functions and variables. Even though we argued national curriculum reforms might have sparked a change, the very meager way in which the Swedish curriculum texts are formulated (Bråting et al., 2019) may mean textbook producers are given little guidance on how to include new aspects of algebra. The low level of national curriculum guidance may also be why the two examined book series can be so different regarding algebra content, particularly in grades 4-6.

Moving on to discuss our framework, let us first discuss our results in the light of similar research conducted. Using a similar page-based analytical method as ours, Bråting and colleagues (2019) examined two Swedish textbook series for grades 1-6, where one was $A_1$ and A for grades 4-6 using the categories generalized arithmetic, equivalence, expressions, equations, and inequalities (EEEI) and functional thinking (Blanton et al., 2018). They found that generalized arithmetic was almost completely missing from $A_1$ and A. The lack of generalized arithmetic is a discrepancy relative to our finding since their category, generalized arithmetic, is, by definition, quite similar to our class *algebra as structure.* Examining this discrepancy is worthwhile since it helps us identify some particularities with our framework and perhaps some affordances. Two reasons explain why Bråting et al. (2019) found so few instances of generalized arithmetic. First, in Bråting et al.'s framework, equivalence, equality, and related issues comprised a separate category which was well represented. In our framework, equality items were classified as *algebra as structure* if they shed light on the meaning of equality and as *algebra as operations on non-numerical symbols* if the focus was on an unknown quantity, like in 3+_=5. Task series like 1+_=10, 2+_=10…, that is, 10-pals, in our classification concerns number composition and decomposition –a structural issue. As we understand, such tasks would end up in the EEEI category in Bråting et al.'s classification. Second, at closer examination, it seems like Bråting et al. put relatively high demands for what counts as generalized arithmetic and that their category of generalized arithmetic most closely resembles the formal or perhaps even the explicit variant of the structure class in our framework, which also in our analysis was not very frequent. Our framework is also generous in including items in the potential structure category.

Why do we feel our framework is worthwhile exploring instead of, for example, following Blanton et al.'s (2018) lead like Bråting and colleagues (2019)? Firstly, as explained in our framework section, we think that operating on unknowns, letters, and other non-numerical symbols in various forms in expressions that may or may not contain the equal sign using the knowledge you have from operating on numbers and operations deserves a separate class, as this is the most classic form of school algebra. When you allow the *potential* variant of this algebra type, our analysis shows it can be introduced relatively early.

Secondly, and for us, more importantly, the algebra type we call *arithmetic structure* is something else. Here, the intention is to use symbols and other means for making general relationships among arithmetic objects, operations, and relations visible. Algebraic techniques are used to shed light back on arithmetic. Our analysis shows it is pretty easy to introduce potential structural algebra elements but harder to follow up on the formal and explicit levels. We believe working more on explicating arithmetic structure would be worthwhile for textbook procedures.

The research presented here is the first time we use the framework and should be considered somewhat preliminary. There might be cases where our classifications are not consistent, and the definition of the four classes and the explicitness levels can probably improve. We still submit that even the current version of the framework was functional in providing insights into essential differences between the two book series and trends manifested in both series. While school algebra and early algebra, is already well-researched, we hope our framework can help clarify some important issues.

## REFERENCES

Blanton, M., Brizuela, B. M., Stephens, A., Knuth, E., Isler, I., Gardiner, A. M., Stroud, R., Fonger, N. L., & Stylianou, D. (2018). Implementing a Framework for Early Algebra. In C. Kieran (Ed.), *Teaching and Learning Algebraic Thinking with 5- to 12-Year-Olds* (pp. 27–49). Springer International Publishing.

Bråting, K., Madej, L., & Hemmi, K. (2019). Development of algebraic thinking: Opportunities offered by the Swedish curriculum and elementary mathematics textbooks. *Nordisk Matematikkdidaktikk, NOMAD:[Nordic Studies in Mathematics Education]*, *24*(1), 27–49.

Freudenthal, H. (1983). *Didactical phenomenology of mathematical structures*. Reidel.

Kaput, J. J. (2008). What Is Algebra? What Is Algebraic Reasoning? In J. J. Kaput, D. W. Carraher, & M. L. Blanton (Eds.), *Algebra in the early grades* (pp. 5–18). Lawrence Erlbaum Associate.

Kieran, C. (2004). Algebraic thinking in the early grades: What is it. *The Mathematics Educator*, *8*(1), 139–151.

Kieran, C. (2018). Seeking, Using, and Expressing Structure in Numbers and Numerical Operations: A Fundamental Path to Developing Early Algebraic Thinking. In C. Kieran (Ed.), *Teaching and Learning Algebraic Thinking with 5- to 12-Year-Olds* (pp. 79–105). Springer International Publishing.

Kieran, C., Pang, J., Schifter, D., & Ng, S. F. (2016). *Early algebra: Research into its nature, its learning, its teaching*. Springer Nature.

Mason, J. (2018). How Early Is Too Early for Thinking Algebraically? In C. Kieran (Ed.), *Teaching and Learning Algebraic Thinking with 5- to 12-Year-Olds* (pp. 329–350). Springer.

Mulligan, J., Mitchelmore, M., & Prescott, A. E. (2006). Integrating concepts and processes in early mathematics: The Australian Pattern and Structure Awareness Project (PASMAP). In J. Novotná, H. Moraová, M. Krátká, & N. Stehlíková (Eds.), *Proc. 30th Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 4, pp. 209–216). PME.

Radford, L. (2022). Introducing equations in early algebra. *ZDM–Mathematics Education*, *54*(6), 1151–1167.

# INSTRUCTIONS IN MATH PROBLEMS: ARE PROOF TASKS CONSIDERED MORE DIFFICULT BY UNIVERSITY STUDENTS?

Lukas Hellwig, Sebastian Geisler

University of Potsdam (Germany)

*Proof tasks are the most used tasks in university mathematics programs. They pose a particular challenge at the transition from school to university because first-year students often have little experience with proofs from school. It can therefore be assumed that they are less likely to tackle these tasks or to find them easy. This study examines the extent to which students are guided by the wording – in terms of used operators – of proof tasks when assessing them. N=298 first-year students were surveyed. A MANOVA revealed no significant differences in situational interest, self-efficacy and perceived difficulty that could be attributed to the used operators. However, personal characteristics had an influence on the perceived difficulty of tasks. Implications of the results for further research are deduced and discussed.*

## INTRODUCTION

Mathematics undergraduate students spend their learning time mainly in lectures or solving exercise tasks (Rach et al., 2014). Both learning situations are therefore likely to have significant influence on the learning process of students. While students are guided by lecturers in courses, they usually work on exercise tasks unsupervised by professionals (Rach et al., 2014). This is problematic at the transition from school to university because not only the subject matter of university mathematics changes due to its focus on strict proof and formal language, but students are also expected to develop into autonomous learners (Gueudet, 2008). While being the most common task type at university (Weber & Lindmeier, 2020), proof tasks are underrepresented in German textbooks at school (Vollstedt et al., 2014). So first-year students are not familiar with proof tasks at the beginning of their studies and have problems with solving them (Rach et al., 2014). These results are consistent with international studies (Stylianou et al., 2015). As an unfamiliar task format, proof tasks could therefore be perceived challenging by first-year students. This study therefore examines whether such tasks are rated as difficult, intimidating or uninteresting and which influence the wording of the task (in particular the used operator) has on students' perception.

## THEORETICAL BACKGROUND

### Exercise tasks at university and task difficulty

The difference between university and school mathematics is also reflected in the type of exercise tasks students are given. Recent studies distinguish between the type of proof tasks and the type of arithmetic tasks at university (Weber & Lindmeier, 2020). Arithmetic tasks can be solved with only schematic internal mathematical processes, while proof tasks require formal deductive reasoning (Rach et al., 2014). The latter



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 17-24). PME.

make up the majority of the tasks mathematics majors and pre-service teachers are expected to complete (Weber & Lindmeier, 2020). To specify the level of expectation, tasks usually include a prompt (e.g., “Prove”), which is later on referred to as their operator.

A distinction must be made between confirmatory and exploratory proof tasks. Confirmatory tasks (e.g., using the operator “Prove”) reveal that the assertion in question is actually true. Exploratory tasks (e.g., using the operator “Investigate”) leave this question open, so that students must first make an assumption and then prove it. As Neubrand et al. (2002) note, the openness of a task (synonymous with the existence of several possible solutions) is also a factor for its difficulty.

Proof tasks are particularly challenging for students (Weber, 2001), which can be problematic, because the individually perceived task difficulty is very important for the personal decision to tackle a task or not (Street et al., 2022). This perception can differ between individuals, for example if students can remember comparable tasks (Street et al., 2017). In this respect, students can “get used to” difficult tasks and then rate them as less difficult than inexperienced students. It is currently still relatively unclear which criteria contribute significantly to this perception.

**Self-efficacy, interest and achievement**

The scope of one’s abilities could influence the aforementioned perception. A general view of one’s own abilities in a particular subject is referred to as one’s Self-concept regarding that subject (Bandura, 1986). It is considered independently of the specific content or learning situation and stable over time (Bandura, 1986). Self-concept can be differentiated according to different facets of mathematics like self-concept with regard to school or university mathematics (Rach et al., 2019). Self-efficacy, on the other hand, refers to one's own perception of the abilities required to achieve a certain performance (Bandura, 1986) and is therefore a content- or task-related and temporally variable state (Bandura, 1997). Students with high self-efficacy will complete a difficult task with greater persistence, effort and accuracy, because they believe they have the necessary skills to solve this task (Bandura, 1986). In this case, they also achieve better learning outcomes, as studies already have shown (Peters, 2013). The reverse also seems to hold true: Good performances in the past influence self-efficacy for future challenges via the experience of mastery (Bandura, 1997). However, these must be actual challenges, as Street et al. (2022) were able to demonstrate this effect only for moderately difficult and difficult, but not for simple tasks.

According to Krapp’s person-object theory, a person’s engagement with a task is further influenced by this person's interest, which can also be subdivided into personal and situational components (Krapp, 2007). Personal interest is a temporally stable, individual variable (trait), while situational interest mostly exists “only for a limited period of time and [is] triggered by external incentives” (Krapp, 2007, p. 7) and is thus considered as a state. Ufer et al. (2017) were able to show that personal interest can be divided into interest in school mathematics and interest in university mathematics.

Previous studies have mostly shown no influence of interest on mathematical performance (e.g., Kosiol et al., 2019). However, interest in university mathematics has a positive influence on study satisfaction and motivation (Kosiol et al., 2019).

## RESEARCH QUESTIONS

As outlined above, the present study focuses on proof tasks and how they are perceived by first-year students in terms of self-efficacy, situational interest and perceived difficulty. In particular we want to answer the following research questions:

Q1: Do the perceptions of tasks that have different operators differ with regard to the task-specific self-efficacy, situational interest and perceived difficulty?

*Theoretically, it should be assumed that exploratory operators are rated as more difficult and that students report lower self-efficacy for them. On the other hand, since confirmatory proof tasks are more uncommon for first-year students, it is plausible that they perceive these operators as more difficult and report lower self-efficacy. This research question is therefore treated as rather exploratively.*

Q2: To what extent do the personality traits (self-concept, personal interest) correlate with the reported perceptions of task specific self-efficacy, situational interest and perceived difficulty?

*It can be assumed that the self-concept correlates positively with the task-specific self-efficacy and negatively with perceived difficulty (H1) while the personal interest will correlate positively with the task-specific situational interest and negatively with perceived difficulty (H2) (Krapp, 2007).*

Q3: Are there differences in these correlations between exploratory and confirmatory tasks?

*As there is not much research concerning the effects of used operators in university tasks, we do not have specific hypotheses and treat this question as explorative.*

## METHODS AND DESIGN

### Selection of operators

In order to identify the operators used for the main study, $N_1 = 117$ tasks from the topic sequences and series – as a canonical topic in real analysis lectures that German undergraduates usually attend in their first term – were examined. The exercises were taken from exercise sheets from five different courses at four public universities in Germany, which were held between the winter term 2022 and 2023.

The evaluation revealed that the three most frequently used operators were “Show” (40 occurrences), “Prove” (22 occurrences) and “Investigate” (16 occurrences). While the first two are classified as confirmatory tasks, the latter belongs to the exploratory tasks. In order to maintain a balance between confirmatory and exploratory tasks and to have a reference task, the open question (the absence of an operator) was included as an exploratory task, which was identified three times in the sample of examined tasks.

**Sample and instruments**

In order to answer the research questions, a questionnaire asking for students' personal interest and self-concept regarding university mathematics as personal traits was used. In addition, a task concerning the convergence of series was presented to the students. This task existed in four versions differing in the aforementioned operators (Prove, Prove, Investigate, Question). The tasks were randomly distributed and identical in content except for their operator. The students were instructed to not work on them. An example task is shown in Fig. 1.

**Exercise 2**
Let $(a_n)_{n\in\mathbb{N}}$ be a real-valued sequence with

$$a_n := \frac{1}{n^2 - n}.$$

Does the series $\sum_{n=2}^{\infty} a_n$ converge?

*Note:* You may use, that $\sum_{n=1}^{\infty} \frac{1}{n^k}$ converges for all real $k > 1$, otherwise diverges.

Fig. 1: Exemplary stimulus with open question

As the students should already have covered the content of the task in their course, the survey had to be carried out halfway through the term. Students were also asked to rate the task-difficulty and were asked about their self-efficacy and situational interest regarding this task. All used scales are shown in Table 1 and were answered on a 5-point Likert scale, with answers ranging from "Does not apply" (1) to "Applies" (5).

| Scale | Items | Reliability | Example Item | Source |
|---|---|---|---|---|
| Personal interest | 5 | .93 | I am interested in the kind of mathematics that I learn at university. | Rach et al. (2021) |
| Self-concept | 3 | .84 | The math that is practiced at university is easy for me. | Ufer et al. (2017) |
| Self-efficacy | 1 | | I am confident that I can solve this task. | Adapted from (Willems, 2011) |
| Situational interest | 1 | | I am interested in this task. | Adapted from (Willems, 2011) |
| Perceived difficulty | 1 | | I find this task difficult. | Own development |

Tab. 1: Number of Items, Reliability Scores (Cronbachs $\alpha$), Source and example Items for used scales

The sample for answering question 1 consists of $N_2 = 298$ students (42% female) from four public German universities. The data was collected in six "Real Analysis" courses in the summer and winter terms of 2023, which were aimed at mathematics majors and pre-service teachers (67% pre-service teachers, 91% first-year students). The other two research questions are answered with a reduced sub-sample of $N_3 = 136$ students (40% female, 52% pre-service teachers, 83% first-year students), as personal interest and self-concept were not collected in one course.

## RESULTS

To answer the first research question, a MANOVA was conducted which found no relationship between the task's operator and task-related self-efficacy, situational interest and task difficulty ($F(9{,}710.8) = 0.351, p = 0.957, \eta^2 = 0.004$). However, there are small descriptive differences between the operators, as shown in Fig. 2.

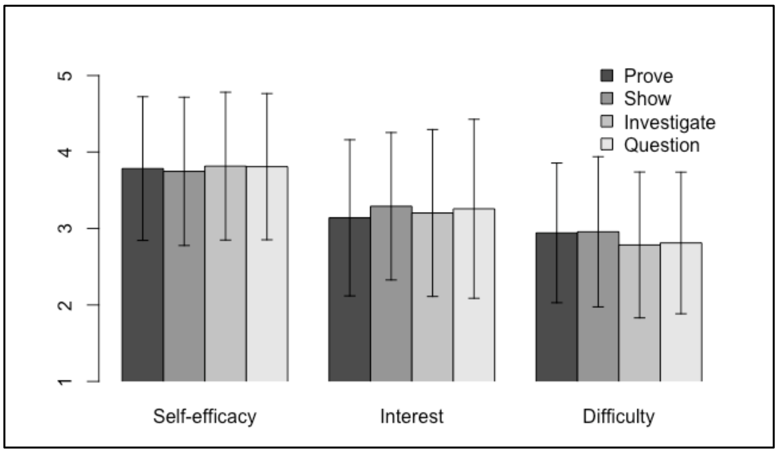



Fig. 2: Reported task-related variables (means and standard deviations) by operator, answers between *1="Does not apply"* and *5="Applies"*, $n_{\text{Prove}} = 73, n_{\text{Show}} = 74, n_{\text{Investigate}} = 78, n_{\text{Question}} = 73$

To answer the second research question, a correlation analysis for situational state variables (self-efficacy, situational interest and perceived difficulty) and personal trait variables (personal interest, self-concept) in the entire $N_3$ subsample was calculated. The results are shown in Table 2. The personal interest shows the strongest positive correlation with the task specific interest, confirming H1, and is least correlated with the perceived difficulty. The self-concept is relatively equal correlated with all reported state variables and thus fairly positively correlated to the task-related self-efficacy and negatively with the perceived difficulty. This confirms H2.

| | Self-efficacy | Situational interest | Perceived difficulty |
|---|---|---|---|
| Personal interest | .382 | .555 | -.182 |
| Self-concept | .389 | .324 | -.311 |

Tab. 2: Correlation Analysis (Pearson-Coefficient) for state and trait variables, all results are significant with $p < 0.01$, $N_3 = 136$

Research question 3 was investigated using multiple correlation analyses for state and trait variables in the $N_3$ subgroups divided by the operators. The analyses results are presented in Tab. 3. Personal interest shows a medium to strong positive correlation with situational interest across the operator subgroups (while being not significant in the subgroup of the “Show”-Operator). Self-concept is significantly positive correlated with the self-efficacy in all subgroups. Interestingly the negative correlation of perceived difficulty and self-concept or personal interest vanishes (not significant) in the explorative operators’ subgroups.

| Correlation | Prove $n = 33$ | Show $n = 35$ | Investigate $n = 36$ | Question $n = 31$ |
|---|---|---|---|---|
| Personal interest & self-efficacy | .453** | .331 | .377* | .369* |
| Personal interest & situational interest | .735** | .310 | .546** | .469** |
| Personal interest & perceived difficulty | -.306 | -.226 | -.006 | .010 |
| Self-concept & self-efficacy | .439* | .389* | .342* | .354* |
| Self-concept & situational interest | .382* | .326 | .164 | .135 |
| Self-concept & perceived-difficulty | -.568** | -.325 | .034 | .014 |

Tab. 3: Intragroup-correlations (Pearson-Coefficient) for trait and state variables grouped by operator, $*p < 0.05$, $**p < 0.01$

**DISCUSSION**

Results show that no difference could be found between the operators with regard to students’ perception of tasks. Neither did students report differences influenced by the operator in terms of self-efficacy or perceived task difficulty, nor in terms of situational interest. This could be explained by the fact that students had already developed a sense for task formulations. The task versions given were all proof tasks in terms of the type of mathematical work (Weber & Lindmeier, 2020). Results suggest that students recognized this despite the various operators and evaluated the tasks accordingly.

However, the results also show that personal characteristics are more important for students' perception of the tasks than the used operator. A clear relation between state and trait variables is not surprising, because trait variables are considered to influence students’ state variables (e.g., Krapp, 2007). It is noteworthy that in our sample, personal interest correlates about as strongly with self-efficacy as self-concept, which is consistent with the results of Nuutila et al. (2020).

Interestingly, the results for research question 3 reveal that the relations between trait and state variables are not independent of the used operators. First of all, it should be noted that a high level of interest in university mathematics predicts a high task-related interest in a task with “Prove” operator better than with exploratory operators. So, the personal and situational interest tend to coincide in these cases. One could therefore assume, that these tasks are more in line with the expectations of students regarding university mathematics, which fits to the results of Rach et al. (2014). Similar results can be seen for the self-concept of university mathematics and the reported perceived task difficulty of "Prove" tasks. However, the correlations for the "Prove" operator and the "Show" operator sometimes differ more than for the "Prove" operator and both exploratory operators. Thus, the theoretical assumption of a coherent type of confirmatory operators does not appear to be suitable for interpreting these results with regard to students' perception.

## Limitations and outlook

The survey of the study was conducted with an authentic, subject-related challenging stimulus and during the lecture course, so that not only positive selection due to students dropping out distorts the sample, but also a habituation effect among the students to the task formats (and formulations) of their lecturer could have influenced the answers. Additionally, our study relied on self-reported state variables, which makes our data vulnerable to over- or underestimation. In general, the results presented for research questions 2 and 3 should be interpreted with caution, as the sample size of the present study is rather small. Nevertheless, it does raise some interesting questions.

Results indicate that the openness of a proof task only leads to minimal differences in students’ perception of the task – contrary to the results of Neubrand et al. (2002). Thus, it is not only necessary to clarify the degree to which students perceive this openness, but also how important they perceive it for their judgment. Moreover, the question arises what differences students see concerning the level of expectations implicitly formulated by the operators used in tasks. As students do not seem to make any differences in their assessment of difficulty and interest, one could wonder what point having nuanced instructions in tasks makes and to what extent they are (and can be) used by lecturers in a goal-oriented manner.

Our ongoing research will now focus on other characteristics of tasks (like the complexity of necessary concepts to solve them) to explain differences in students’ situational interest, self-efficacy and perceived difficulty.

## References

Bandura, A. (1986). *Social foundations of thought and action: a social cognitive theory*. Prentice-Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W.H. Freeman and Company.

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

Kosiol, T., Rach, S., & Ufer, S. (2019). (Which) Mathematics Interest is Important for a Successful Transition to a University Study Program? *International Journal of Science and Mathematics Education*, 17(7), 1359–1380.

Krapp, A. (2007). An educational-psychological conceptualisation of interest. *International Journal for Educational and Vocational Guidance*, 7(1), 5–21.

Neubrand, M., Klieme, E., Lüdtke, O., & Neubrand, J. (2002). Kompetenzstufen und Schwierigkeitsmodelle für den PISA-Test zur mathematischen Grundbildung. *Unterrichtswissenschaft*, 20(2), 100–119.

Nuutila, K., Tapola, A., Tuominen, H., Kupiainen, S., Pásztor, A., & Niemivirta, M. (2020). Reciprocal Predictions Between Interest, Self-Efficacy, and Performance During a Task. *Frontiers in Education*, 5(36).

Peters, M. L. (2013). Examining the Relationships among Classroom Climate, Self-Efficacy, and Achievement in Undergraduate Mathematics: A Multi-Level Analysis. *International Journal of Science and Mathematics Education*, 11, 459–480.

Rach, S., Heinze, A., & Ufer, S. (2014). Which Mathematical Demands do Students Expect in the First Semester of Their Study? *Journal Für Mathematik-Didaktik*, 35(2), 205–228.

Rach, S., Ufer, S., & Kosiol, T. (2019, February). Self-concept in university mathematics courses. *Eleventh Congress of the European Society for Research in Mathematics Education.*

Rach, S., Ufer, S., & Kosiol, T. (2021). The role of self-concept when studying mathematics — Do students feel fit in mathematics? *Zeitschrift Für Erziehungswissenschaft*, 24(6), 1549–1571.

Street, K. E. S., Malmberg, L. E., & Stylianides, G. J. (2017). Level, strength, and facet-specific self-efficacy in mathematics test performance. *ZDM*, 49(3), 379–395.

Street, K. E. S., Stylianides, G. J., & Malmberg, L. E. (2022). Differential relationships between mathematics self-efficacy and national test performance according to perceived task difficulty. *Assessment in Education: Principles, Policy and Practice*, 29(3), 288–309.

Stylianou, D. A., Blanton, M. L., & Rotou, O. (2015). Undergraduate Students' Understanding of Proof: Relationships Between Proof Conceptions, Beliefs, and Classroom Experiences with Learning Proof. *International Journal of Research in Undergraduate Mathematics Education*, 1(1), 91–134.

Ufer, S., Rach, S., & Kosiol, T. (2017). Interest in mathematics = interest in mathematics? What general measures of interest reflect when the object of interest changes. *ZDM*, 49(3), 397–409.

Vollstedt, M., Heinze, A., Gojdka, K., & Rach, S. (2014). Framework for examining the transformation of mathematics and mathematics learning in the transition from school to university: An analysis of german textbooks from upper secondary school and the first semester. In S. Rezat, A. Peter-Koop, & M. Hattermann (Eds.), *Transformation - A Fundamental Idea of Mathematics Education* (pp. 29–50). Springer.

Weber, B.-J., & Lindmeier, A. (2020). Viel Beweisen, kaum Rechnen? Gestaltungsmerkmale mathematischer Übungsaufgaben im Studium. *Mathematische Semesterberichte*, 67(2), 263–284.

Weber, K. (2001). Student difficulty in constructing proofs: The need for strategic knowledge. *Educational Studies in Mathematics*, 48, 101–119.

Willems, A. S. (2011). *Bedingungen des situationalen Interesses im Mathematikunterricht*. Springer.

# ILLUSTRATING A METHOD FOR ANALYZING MULTIMODAL ARTIFACTS USED IN TRANSACTIONS OF PRACTICE

Patricio Herbst, Gil Schwarts, and Amanda Brown

University of Michigan, USA

*We illustrate how concepts from systemic functional linguistics are adapted for the analysis of multimodal representations of practice used in activities where teachers and teacher educators transact meanings about practice. We focus on the transactive register used to project practice meanings to the audience of these representations. We showcase the systems called visibility (how much of the classroom experience happening is made visible to the viewer), temporality (how sequence and duration of events are represented), and theme (how semiotic resources maintain and develop themes). We apply these systems to examine the differences between two storyboards of algebra lessons that were used in a professional development context and the different kinds of reactions teachers offered to the different storyboards.*

## INTRODUCTION

We contribute to the examination of teachers' learning with multimodal representations of practice (RoP). Teacher educators (TE) have long been using RoP, in the form of various media types (e.g., written cases, videos, simulations, transcripts, animations, storyboards) to engage (prospective or practicing) teachers in activities where they can learn in, from, and for practice (Lampert, 2010). Whilst RoP might be differentiated by their media types, they also have common characteristics that issue from their multimodality, purpose, and subject matter. We offer a multisemiotic analysis of RoP that especially accounts for how they support transactions of practice between TE and teachers. We tackle the question: When teachers annotate a multimodal RoP, what aspects of the RoP need to be analyzed apriori to make sense of their comments? We illustrate our analytic approach by focusing on storyboards designed with cartoon characters and raise the question of how different semiotic choices relate to how practicing algebra teachers engage with the represented practice.

## THEORETICAL FRAMEWORK

This paper provides empirical illustration of the theoretical contribution by Herbst, Chazan, and Schleppegrell (2023), which draws on systemic functional linguistics (SFL; Halliday & Matthiessen, 2004) to organize the multimodal resources available to producer and consumer of a storyboard to construe meaning. While SFL was originally developed for language, various extensions to other semiotic systems (e.g., displayed art, children's picture books, film; O'Toole, 2011; Painter et al., 2013; O'Halloran, 2004) have encouraged us to extend it to the case of RoP. Storyboards are sequences of frames, each of which includes graphics and written language, which are commonly used in the design of animations, films, and graphic novels; they have been used in teacher education and in research on teaching for more than a decade and



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 25-32). PME.

sometimes called comics, vignettes, or scenarios (Friesen & Knox, 2022; Herbst et al., 2011; Lin, 2023). Like the children's picture books analysed by Painter et al. (2013), the combination of images and writing make storyboards a multimodality with which the multimodality of classrooms (i.e., the use of oral and written language, gesture, body language, and facial expression) can be represented and transacted. Herbst et al. (2023) used SFL's metafunction dimension to organize the search for and identify some of the systems available in the storyboard modality to make meaning.

The metafunction dimension of SFL proposes that multimodal texts in a semiotic system fulfill three different metafunctions: (a) ideational – to represent the world and the experiential and logical relationships in the world, (b) interpersonal – to relate transaction partners, particularly producer and consumer, and (c) textual – to characterize types of texts. In this paper, we define and illustrate three multimodal systems at play in storyboards of practice—temporality, visibility, and theme—which contribute respectively to the ideational, interpersonal, and textual metafunctions.

SFL's founder M. A. K. Halliday originally proposed the notion of register to describe patterns in the use of elements of language in social context (see Matthiessen et al., 2010, p. 176). Thus, the mathematical register is a variation of language used to make mathematics meanings (where words like center or let are used differently than in everyday English). As SFL has become a social semiotic, capable of analyzing texts in diverse modalities, register has come to be identified with three elements that specify its context: Field (what the communication is about), Tenor (what social relationships are enacted through communication), and Mode (what kinds of texts are used). Christie (2002) contributed to bring SFL closer to education research by identifying two registers in classroom language: The instructional register (used to communicate the content of instruction and apparent, for example, in how a textbook might display the solution of an example problem) and the regulative register (used to organize pedagogy and the classroom experience; Schleppegrell & Oteiza, 2023). Christie characterizes classroom discourse by saying that the regulative register projects the instructional register—patterns of pedagogical language use are used to communicate patterns in mathematical language use. Storyboard RoP are multimodal texts whose ideational metafunction includes representing classroom actors and events. Those RoP used in teacher education or professional development support and help constitute relationships between teacher learners and teacher educators. And they do so in the form of lesson representations to be perused or annotated, lesson plans to be created collaboratively, or exercises in which teachers have to complete or select missing elements of the RoP (Kalinec-Craig et al., 2021; Rougée & Herbst, 2018). Perhaps the case of transcripts is the clearest illustration that representations of practice are not equal to events of classroom practice—a RoP actually uses signs to project the regulative and instructional registers of classroom practice. Herbst et al. (2023) introduced the notion of transactive registers to name patterns of use of semiotic choices to represent classroom practice for transactions among social actors and through particular texts. Though some transactive registers are also in play in instruments designed for research on teaching (e.g., Skilling & Stylianides, 2020), we focus here on the register used in transactions of practice in teacher education.

The transactions of practice that take place in teacher education are more than opportunities to peer over a classroom experience; they rather are pedagogical encounters, and multimodal resources are used in particular ways to enable those RoP to be pedagogical. Thus, the transactive register consists of semiotic systems of choice to construe field, tenor, and mode of situations which constitute the pedagogical relationship among teachers and TEs about practice.

We illustrate how the transactive register allows us to examine records from transactions of practice, specifically two commentaries of algebra lessons. The lessons were represented using storyboards that had similarities and differences. The RoP can and should be examined in terms of what mathematical and pedagogical meanings they construe, but we concentrate here in illustrating the teacher education meanings construed with the transactive register. We look at those in terms of three transactive systems of meanings. One is visibility, which supports the construal of interpersonal meanings (or the tenor of the relationship between producers and viewers) by making classroom experience more or less visible for the participants of the teacher education transaction: For example, in a storyboard, the designer may choose to provide whiteboard content which is interpretable by the viewer (e.g., an inscription in legible mathematical symbols) instead of providing indices that suggest content is on the whiteboard though it cannot be retrieved by the viewer (e.g., a scribble); this suggests that the RoP producer may intend the receiver to read the specific content rather than expect the receiver to bracket it. The second is temporality, which supports the construal of ideational meanings by developing of a sense of sequence and duration of events in the lesson. For example, the left-to-right juxtaposition of frames to indicate before-after sequences of events, or the possibility to insert a frame in between to frames to represent events that avowedly happened between the events represented in the original two frames. The third system is theme, which in the analysis of paragraphs in language refers to progression from the setting of a topic to comment on the topic, to the topicalization of elements of that comment. For the analysis of storyboards, we identify multimodal resources that support a sense of continuity and progression of themes. In this analysis, we identify resources that construe themes across frames. The development of a lesson requires the passing of time and the evolution of discourse. What are the resources that permit the reader to understand new frames as dealing with the same lesson even when elements of prior frames are not present in new frames and what are the resources that permit the reader to identify what is new in each frame?

## MODES OF INQUIRY AND DATA SOURCES

Analyzed data comes from Story*Circles* (Herbst & Milewski, 2018), a professional learning program where teachers participate in online activities of scripting, visualizing, and arguing about a lesson. In this paper, we discuss how RoP were used in two Story*Circles* focused on problem-based algebra lessons that teachers were to represent during a six weeks period. Figure 1 provides the opening and closing frames of the two lessons, which were among the storyboard frames given to participants.

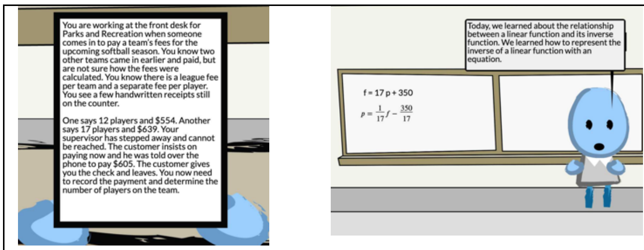



Figure 1a: Bookends of inverse function lesson (L1)

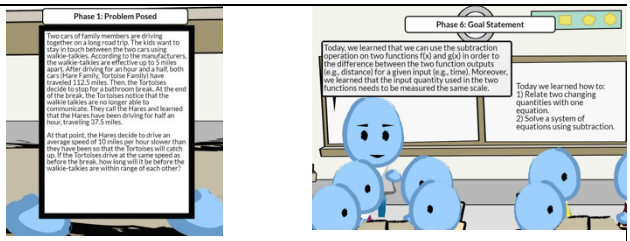



Figure 1b: Bookends of function subtraction lesson (L2)

Each Story*Circle* started with a transaction of practice named "Leave Tracks" where participants (8 and 6 teachers, respectively) were asked to review and annotate a sequence of lesson frames and then answer questions posed, using an annotation application. The design of the two activities differed, as did the ensuing teacher comments; these differences illustrate analysis of the transactive register. We performed a dual analysis of the two activities. An apriori analysis compares the design of the activities in terms of visibility, temporality, and theme. And an aposteriori analysis describes teacher comments in relation to design choices and to program goals.

The RoPs provided to teachers represented lessons which started with a novel mathematical task and ended with the teacher introducing a specific instructional goal as a conclusion of the work on the task. The two ROPs represented similar lesson structure: The teacher launched the task, the students had time to work on it, a whole-class check-in happened after, and the teacher redirected the class to continue working, followed by a whole-class discussion that eventually led to the instructional goal. However, within these similarities the ROPs represented the lessons differently, as we elaborate next. We focus on the moment of the lesson when the teacher redirects the work on the task to make progress toward the lesson goal.

## RESULTS

### Apriori comparison of the two storyboards

We compare the way the lesson was offered to participants in cycles 1 and 2 in terms of *visibility, temporality,* and *theme;* because of space limitations we only share illustrative instances. The inverse function lesson (L1) was represented with a 20-frame storyboard (see Figure 2). In terms of *visibility*, viewers were able to see the teacher, more than a dozen students, the actual writing on the board (in frame 12), and the specifics of what was said. In addition, the text on the board is not hand-written, which could have been a more accurate representation of classroom practice; instead, there was a deliberate choice to make the text readable, at the expense of authenticity. In comparison, in the subtraction of functions lesson (L2) much less is visible: Figure 3b shows a moment of redirection of students' work on the problem in frame 7, and Figure

3a the frame immediately before, but in both cases no board content is to be scrutinized. Furthermore, although the frame zooms-in on the teacher, viewers don't see what's on the board, and they see only about six students, but with very little evidence on what they are doing. We suspect that this low visibility might prompt viewers to imagine, speculate, and eventually script what could have happened. The details visible in L1 might not invite the same behaviors from readers.

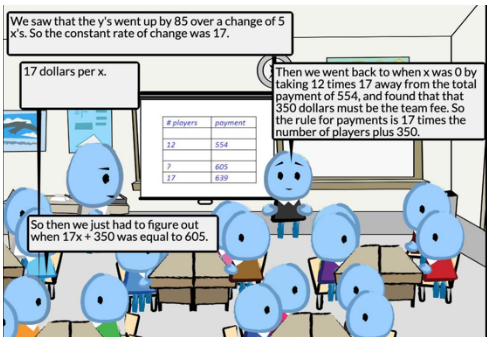



Figure 2a - Frame 12 in L1

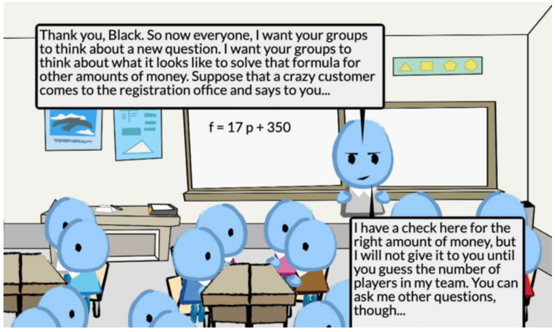



Figure 2b - Frame 13 in Lesson 1

In terms of temporality, in both lessons the serial layout of the frames in the storyboard organizes events in the order in which they happened over time. Yet, they convey different senses of time, using the gutter between frames and ordered actions and speech within each frame. The gutter is used in the comic genre to separate frames and can be used, similarly to a jump cut in video, to fast forward action. The gutter in L1 was used as follows: in frame 13 the words "Thank you, Black" indicate the teacher wrote the equation immediately after the student shown in Frame 12 (wearing a Black vest) presented their solution—signaling that only a few seconds passed between the frames. This sense of temporality may also lead viewers to infer that the teacher was the one who introduced the two variables, including choosing $f$ and $p$ to represent them, since there could not be a major scene omitted in between the frames. In L2, however, the gutter between frames 6 and 7 together with the captions perform a jump cut to a new scene, conveying the sense that significant time has passed between the frames. L1 thus represented a denser temporality (less time between frames), while L2 represented sparser temporality (more time between frames).

To assert that storyboards of practice fulfill a textual metafunction means that each storyboard contains resources that allow it to hang together as one text just as new frames introduce new material. In Figure 2 we note the intersemiotic repetition between Black's statement "17 times the number of players plus 350" in Frame 12 and the inscription "f = 17 p + 350" the teacher made on the board visible on Frame 13. The formula was the comment made by Black on Frame 12 about "the rule for payments" which was the theme. In Frame 13 the formula becomes the theme as the teacher poses a question to the whole class about manipulating the formula. The shot is the same and suggests graphically that the speaker in Frame 12 who was then at the front of the class is back in their seat. That is, the juxtaposed frames not only represent a short time between two events but also permit the analysis of theme development in ways similar as in ordinary discourse analysis (albeit, considering also the repetition and change in graphics). In Figure 3, however, fewer language resources help in the same way. The constancy of the shot and apparent arrangement of the classroom helps suggest these frames are part of the same story. Similarly, the caption is an important resource both for repetition and change. The word "Phase" repeated across frames makes them part of a same story (that one might surmise develops in phases) and the change from 3 in Frame 6 to 4 in Frame 7 suggests that they are consecutive and this enables the inference from the reader that after a discussion of what students had found on the initial problem (presumably Phase 3 of the lesson), the teacher gave them a variation of the problem (announced in Frame 7).

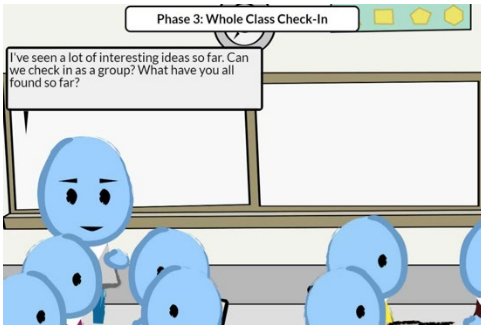



Figure 3a. Frame 6 in L2

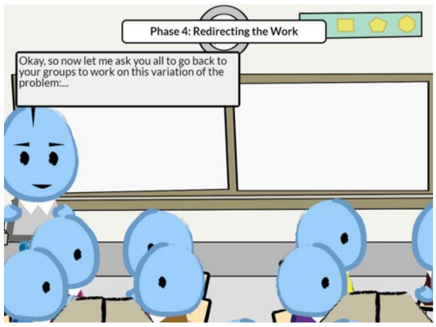



Figure 3b. Frame 7 in L2

**A Posteriori comparison of teacher commentaries**

We present here representative examples of participants' responses to the two activities, focusing on the different ways in which participants described the teacher's moves. Comments on Frame 13 (L1) were characterized by appraising the teacher's move, either negatively or positively. However, this was done in different ways. One participant placed himself as an outsider to the situation:

> I like the extension/generalization to other situations... and bundling it into a "real"-ish story.

Another participant put himself in the scenario, though without agency to change the represented move:

> I'm trying to think how my students would respond here since the teacher is giving them two unknowns now. I can imagine this frustrating some students; which isn't necessarily a bad thing. I think the questions that the students ask on the next few slides kind of display this frustration.

But a third participant put herself in the teacher's shoes and suggested an alternative move:

> I probably would have given each group time to adjust their answers to follow what was shared with the class before moving on to a new problem.

This last comment negatively appraises the represented move, by arguing students needed more time. We suggest this criticism was supported by the temporality choices discussed above, which conveyed the sense that only a few seconds had passed between frame 12 where a student showed their solution, and frame 13.

Overall, we argue that the appraisals shown above are supported by the choices of denser temporality and high visibility of the events in the lesson through the representation that was offered to the participants in L1. The temporality and visibility systems were used to convey the sense that what was visible to the viewers is more or less what happened in the lesson, similarly to when watching videos of practice.

In L2 comments on the redirection moment (L2-7) were different. A first noticeable difference between the cycles is that the participants suggested what could be done:

> Perhaps [the teacher should] give some noticing... [“]I see some great strategies so far that are helping you find out when the vehicles will be within that 5-mile range[”]...[“] it was nice seeing a variety of representations[”]... [“]this group has tried using a graph to represent the situation[”]... [“]this group has started with a table[”].

However, that was not always subject specific, as the following comment shows:

> The teacher could talk with the students about the pros and cons of each of the different ideas that the students have come up with.

This suggests that a sparser representation with no visible student work may not be enough to support participants’ engagement in scripting both the regulative and instructional registers of classroom discourse. L2 had been created purposefully sparser, as our discussion of theme suggests, to engage participants in scripting the various phases of the lesson. But the generic nature of this sparse representation was apparently not sufficient to get participants to specify alternative actions. The facilitator was able to get teachers to act on the suggestions from the comments by bringing in samples of students’ work (as described in Brown et al., 2021).

## CONCLUSION

We illustrated how different choices in the transactive register can be associated with different types of participants interactions with RoP. A denser storyboard (one with denser temporality, more visibility, and subject-specific markers of thematic development) enabled substantive, albeit reactive comments. A sparser storyboard in contrast did enable some suggestions for what to do in between but these suggestions stayed generic during the perusal of the storyboard. Quite often when creating representations of lessons for their use by teachers, teacher educators need to make choices that include selecting, editing, and augmenting media. This suggests that designers can finetune the RoP they present to teachers in anticipation of the kinds of engagement with the lesson they want to enable (e.g., when deciding how to edit and present a video of a lesson). The systems of visibility, temporality, and theme are among the systems that can assist designers in making those design choices.

### Additional information

Research supported by James S. McDonnell Foundation grant 220020524.

## References


Brown, A. M., Stevens, I., Herbst, P., & Huhn, C. (2021). Confronting teachers with contingencies to support their learning about situation-specific pedagogical decisions in an online context. In K. Hollebrands, R. Anderson, & K. Olivier (Eds.), *Online learning in mathematics education* (pp. 291–316). Springer.

Christie, F. (2002). *Classroom discourse analysis: A functional perspective*. Bloomsbury.

Friesen, M., & Knox, A. (2022). Pre-service teachers learn to analyse students' problem-solving strategies with cartoon vignettes. In *Twelfth Congress of the European Society for Research in Mathematics Education* (CERME12). Bolzano-Bozen, Italy. hal-03745715

Halliday, M. A. K. & Matthiessen, C. (2004). *An introduction to functional grammar*. Arnold.

Herbst, P., Chazan, D., Chen, C., Chieu, V.M., and Weiss, M. (2011). Using comics-based representations of teaching, and technology, to bring practice to teacher education courses. *ZDM-Mathematics Education*, *43*(1), 91–103.

Herbst, P., Chazan, D., & Schleppegrell, M. (2023). How digital storyboards support the transaction of practice: The semiotic infrastructure of representations of practice. In B. Pepin et al. (Eds.), *Handbook of digital resources in mathematics education*. Springer.

Herbst, P. and Milewski, A. (2018). What StoryCircles can do for mathematics teaching and teacher education. In R. Zazkis & P. Herbst (Eds.), *Scripting approaches in mathematics education* (pp. 321–364). Springer.

Kalinec-Craig, C. A., Bannister, N., Bowen, D., Jacques, L. A., & Crespo, S. (2021). “It was smart when:” Supporting prospective teachers’ noticing of students’ mathematical strengths. *Journal of Mathematics Teacher Education, 24*, 375–398.

Lampert, M. (2010). Learning teaching in, from, and for practice: What do we mean? *Journal of Teacher Education, 61*(1-2), 21–34.

Lin, Y. C. (2023). Using virtual classroom simulations in a mathematics methods course to develop pre-service primary mathematics teachers' noticing skills. *British Journal of Educational Technology, 54*(3), 734–753.

Matthiessen, C., Teruya, K., & Lam, M. (2010). *Key terms in systemic functional linguistics*. A&C Black.

O’Halloran, K. L. (2004). Visual semiosis in film. In K. O’Halloran (Ed.), *Multimodal discourse analysis: Systemic functional perspectives* (pp.109–130). Continuum.

O'Toole, M. (2011). *The language of displayed art* (2nd ed.). Routledge.

Painter, C., Martin, J., & Unsworth, L. (2013). *Reading visual narratives: Image analysis of children's picture books*. Equinox.

Schleppegrell, M. J., & Oteíza, T. (2023). Systemic functional linguistics: Exploring meaning in language. In M. Hanford and J. P. Gee (Eds.), *The Routledge handbook of discourse analysis* (2nd ed., pp. 156–169). Routledge.

Skilling, K., & Stylianides, G. J. (2020). Using vignettes in educational research: a framework for vignette construction. *International Journal of Research & Method in Education, 43*(5), 541–556.

# CHILDREN'S BEGINNING USE OF MULTIPLICATION IN EARLY PROPORTIONAL REASONING: EXAMINATION OF WRITTEN WORK BY SECOND GRADERS

Keiko Hino[1], Hisae Kato[2], and Hiraku Ichikawa[3]

[1]Utsunomiya University, [2]Hyogo University of Teacher Education, [3]Miyagi University of Education

*In this study, we explore how learning whole-number multiplication relates to progress in early proportional reasoning. We conducted two written surveys of 64 Japanese second-grade children, aged 7−8, before and after learning whole-number multiplication. The change in children's performance depended on the numerical features of the presented problems. We analyzed how they used multiplication to solve the problems in the "after learning" survey and identified four codes on their uses of multiplication: "use the form of the expression," "use in the process of calculation," "use to simplify the problem," and "use to find the relationship between two quantities." We discuss how these codes relate to the change in children's proportional reasoning that they had previously developed.*

## INTRODUCTION

Proportional reasoning refers to "detecting, expressing, analyzing, explaining, and providing evidence in support of assertions about, proportional relationships" (Lamon, 2005, p. 4). Because of its educational importance, research on proportional reasoning has a long history (e.g., Harel & Confrey, 1994; Lamon, 2005; Lobato et al., 2010). However, there is still a question of whether it is adequately fostered through schooling. For example, in the National Assessment of Academic Ability conducted for sixth-grade students in Japan, difficulties with the concepts of multiplication and division of decimals and fractions, ratios, and proportions have been repeatedly pointed out, and issues are raised in the children's qualities and abilities related to handling proportional relationships (e.g., MEXT, 2012, 2018).

In this study, we approach the above issues by focusing on the formation of a foundation for proportional reasoning from the lower grades of primary school (e.g., Hino et al., 2022). Particularly, the lower grades introduce multiplicative concepts that are significant for its formation (Lamon, 2005). How do these concepts contribute to the development of children's proportional reasoning? Models of the development of proportional reasoning are often built on the results of children's surveys, but the relationship with the mathematical content that children are learning at that time is rarely discussed.

Therefore, by focusing on multiplication, we explore how the learning of whole-number multiplication relates to children's progress in early proportional reasoning (EPR) based on written surveys. We address two research questions: (i) what proportional reasoning performance do children exhibit before and after learning



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 9-16). PME.

whole-number multiplication and (i) in what way do children begin to use the learned multiplication in solving proportion problems?

# PERSPECTIVES

## Children's EPR

To capture children's EPR, we used several perspectives from previous research. First, Lobato et al. (2010) listed 10 *essential understandings* of ratio, proportion, and proportional reasoning. Number one on the list is that "[r]easoning with ratios involves attending to and coordinating two quantities" (p. 15). Following this, we examine how children attend to two quantities, as children tend to reason with a single quantity before reasoning with ratios (p. 15).

Second, we examine how children coordinate two quantities using the construct of *a composed unit*. According to Lobato et al. (2010), a ratio as a composed unit is formed by combining two quantities to create a new unit, e.g., "10 cm in 4 seconds." At the beginning levels of proportional reasoning, they listed that "equivalent ratios can be created by iterating and/or partitioning a composed unit" (p. 36). Similar constructs have been employed in other previous studies on upper-grade students (e.g., Kaput & West, 1994; Sigh, 2000). Recently, Vanluydt et al. (2020) conducted a cross-sectional study of EPR in children aged 5–9 considering missing-value problems, which include problems with one-to-many correspondence (e.g., 1p: 2g → 4p: ?g) and many-to-many correspondence (e.g., 2p: 4g → 6p: ?g ). The clusters they found indicated children's progression from one-to-many-to-many-to-many correspondence. Because the many-to-many correspondence problem requires the creation of a composed unit, the result showed progress in the ability of children of this age range to handle composed units. However, there were large individual differences, indicating that the children had not yet fully mastered these skills, even at age 9.

## Whole-number multiplication in the second-grade curriculum in Japan

We also consider how multiplication, which children are introduced to in mathematics lessons, influences children's problem-solving strategies. Whole-number multiplication is introduced in the second grade in the Japanese Course of Study (MEXT, 2018). The aims of this mathematical content include understanding the meaning of multiplication, constructing a multiplication table, and finding and using the properties of multiplication. For the meaning of multiplication, for example, "the total number of oranges when there are 4 plates with 5 oranges in each plate" is expressed as 5 × 4; namely, the equal group structure is taught. Multiplication is considered a concise expression of repeated addition; it is captured as the expression "(unit quantity) × (number of units)." The meaning of "how many times as big as" is also taught. For example, "the length of 3 times as long as 2 meters" is expressed as 2 × 3. This meaning conveys the idea of "measurement" (Isoda & Olfos, 2020, p. 43). Subsequently, in teaching the multiplication table, an array diagram is introduced to be used for contriving effective strategies or checking the results of calculations.

## METHOD

We conducted two written surveys on proportional reasoning in 2022 before (first semester) and after (third semester) the learning of whole-number multiplication in three second-grade classrooms in Japan. In the before-learning (BL) and after learning (AL) surveys, respectively, 64 and 63 responses were obtained. All the children were taught whole-number multiplication in the second semester using textbooks approved by MEXT (2018).

The same problems (P1–P5) were used in the two surveys to determine the price of eggs in a shopping situation (see Table 1). All the problems are missing-value types with discrete quantities, and the numerical feature involves variation in their divisibility. P1 contains one-to-many correspondence (a unit ratio), and P2–P5 contain many-to-many correspondence. Particularly, for P2 and P3, the ratio in the problem, 6 eggs:30 yen, is divisible to obtain the price per egg; meanwhile, it is not for P4 and P5. For P2 and P4, the ratio in the problem can be iterated to find the missing value (18 and 72 eggs, respectively); however, not for P3 and P5. The problems include the word "pack" and illustrations to assist the children in creating composed units.

In the analysis, children's responses to each problem in the two surveys were scored and their strategies were recorded. Moreover, to determine their differences in BL and AL, we examined 60 children who responded to both surveys. First, we classified their BL performance on P1–P5 into six correct-and-incorrect answer patterns: "00000," "10000," "11000," "11100," "11110," and "others," where "1" and "0" indicate correct and incorrect answer, respectively. We created profiles of 47 children belonging to the first five patterns in their changes of responses, including their strategies of using multiplication in solving the problems. Six children had already used multiplication in BL; meanwhile, for any of the problems in the two surveys, four children did not use multiplication. Thus, we excluded these 10 children and analyzed 37 children's responses to produce the results of the use of multiplication reported here. The earlier stage of analysis was performed with school teachers who also helped us during data collection (Hino et al., 2023). Concerning the use of multiplication, initial codes were proposed by the first author of this paper and then independently checked by the other two authors. We have analyzed discrepancies in our coding and revised the codes.

## RESULTS

### Change in the overall performance of children in the BL and AL surveys

Table 1 lists the percentages of correct answers in the BL and AL surveys. The percentages of correct answers to P1 largely increased from BL to AL. For P2–P5, an increase in the percentage of correct answers can be seen in P2 and P4, which are problems that require an answer by recognizing and iterating the composed unit from the problem text. Meanwhile, there was almost no change or decrease in the percentages of correct answers for P3 and P5. Because these problems cannot be solved only with the ratio that appears in the text, the children are required to create a new composed unit by partitioning it by themselves. The children continued to have difficulty handling such fractional parts.

Table 1: Percentages of correct answers for the five written problems

| Problem | BL survey (n = 64) | AL survey (n = 63) |
|---|---|---|
| P1. In a shop, they sell eggs. One egg is 5 yen. You want to buy 7 eggs. What is the price? | 69% | 95% |
| P2. One pack contains 6 eggs and 30 yen. You want to buy 18 eggs. What is the price? | 36% | 44% |
| P3. One pack contains 6 eggs and 30 yen. You want to buy 10 eggs. What is the price? | 27% | 20% |
| P4. In another shop, one pack contains 8 eggs and 60 yen. You want to buy 72 eggs. What is the price? | 9% | 20% |
| P5. One pack contains 8 eggs and 60 yen. You want to buy 20 eggs. What is the price? | 0% | 3% |

For the children's solutions, we only highlight major observations in their change from BL to AL surveys. First, the most obvious change was seen in P1, where children in the BL answered by adding 5 seven times (repeated addition) or counting by 5s, but almost all children in the AL answered by only multiplying or by multiplying and adding 5s in two ways. Second, from P2 to P5, there was a marked increase in the number of solutions that overlooked the "a" in "a eggs and b yen" in the problem text. For example, in P2, some children added 30 18 times or wrote 30 × 18; such solutions were seen in 8 and 19 children in BL and AL, respectively. Third, one of the typical thinking strategies of proportional reasoning, unit rate, uses the price per egg; it decreased from 8 to 2 in P2 and from 6 to 4 in P3 after multiplication learning. Another thinking strategy, building-up, uses the composed unit "up and down" (Lamon, 2005). It was used frequently in P2 and P4 in BL, but its usage decreased from 22 to 21 in P2 and from 20 to 11 in P4. Incorrect or incomplete usage of the building-up strategy was noticeable in BL, but such responses were less salient in AL.

**Children's use of multiplication in the AL survey**

In this section, we describe the results of the analysis of 37 children who belonged to the five correct-and-incorrect answer patterns in BL. We developed four codes for the use of multiplication in AL. Table 2 shows the number of children that are assigned to each code. Below, we illustrate each code in the children's drawn responses to P2–P5.

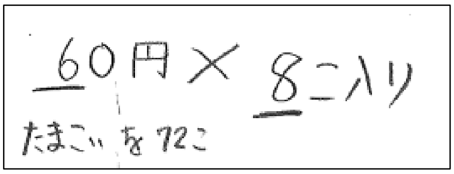
**Use the form of the expression (Form).** Some children's use of multiplication is not based on the meaning that the whole is a repeated addition of a quantity, but only means

Table 2: Number of children in each correct-and-incorrect answer pattern in BL who used multiplication for P2–P5 in AL.

| Code | Number of children | | | | |
|---|---|---|---|---|---|
| | 00000 (n = 13) | 10000 (n = 9) | 11000 (n = 9) | 11100 (n = 5) | 11110 (n = 1) |
| Form | 4 | 1 | 2 | 0 | 0 |
| Simplification | 7 | 6 | 1 | 0 | 0 |
| Calculation | 2 | 4 | 4 | 2 | 0 |
| Relation | 2 | 2 | 5 | 3 | 1 |

Note: when multiple codes are assigned to one child, they are all counted.

a whole that comes up by putting two pieces of information into "a" and "b" in "a × b." For example, a child, G1 (11000), drew the following multiplication in P4 (Fig. 1). G1 also wrote, "The reason is that a pack of 8 pieces costs 60 yen, so I thought I could get the price by multiplication." In Fig. 1, "8" and "60" are underlined. Although the expression is inappropriate, the reason it is written shows that the expression is not based on the meaning of repeated addition. Notably, in BL, G1 understood the problem situation and attempted to capture the relationship among quantities (she was successful for P2). However, in AL, her thinking strategy regressed to a combination of numbers in which multiplication was used as a "form" not only in P4 but also in P3 and even P1.




60 yen × 8 pieces contained

Egg 72 pieces

Fig. 1: G1's drawn multiplication for P4 in AL

**Use to simplify the problem (Simplification).** The problem statement says "b eggs and a yen," but some children missed the condition "b eggs." They used multiplication to find the entire price by denoting "(a yen) × (number of eggs bought)." The problem has been simplified to a multiplication of "a yen for 1 egg" or a multiplication of the meaning of buying several packs. As mentioned above, the use of this type has increased. When looking at the correct-and-incorrect answer pattern in BL, we know that most of those children who were using this meaning of multiplication belonged to the 00000 or 10000 patterns (see Table 2).

**Use in the process of calculation (Calculation)**. The children also used multiplication in their calculation procedures to derive answers. The children's responses included "divide and compute using ×10," "remove 0, compute using multiplication table, and add 0 again," and "use the distributive law to make the calculation easier." For example, Fig. 2 shows the case of a child, H5 (10000), in P3. H5 simplified the problem and wrote "60 × 5 = 300" to show how she derived the answer. In the callout, she expressed both her calculation procedure and its benefits. As in H5's case, children who perceived the problem in a simplified manner were faced with the necessity of adding numbers many times. In an attempt to reduce the calculation burden, multiplication was used as part of the calculation procedure.

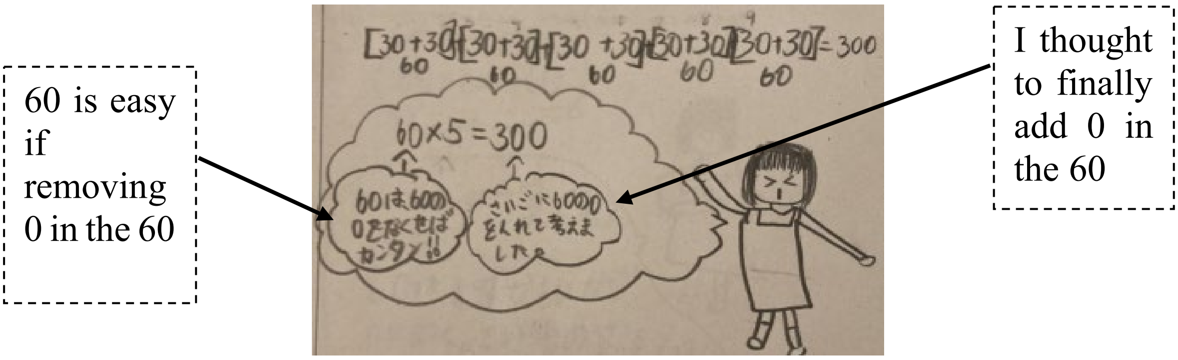



Fig. 2: H5's drawn response to P3 in AL

**Use to find the relationship between two quantities (Relation).** In our final code, multiplication is used to find relationships between quantities, including the relationship between the number of eggs in the problem statement, the relationship between the price in the problem statement and the price to be found, and the relationship between the price of one egg and the price in the problem statement. For example, in P2, the relationship between the number of eggs written in the problem is captured by multiplication (6 × 3 = 18). Meanwhile, when the relationship between the price in the problem and the price to be found is captured by multiplication (30 × 3), "3" is derived from the relationship between the number of eggs. Further, one child, F25 (11000), used multiplication to determine the relationship between the price of one egg and the price in the problem statement, but she was confused about what she got. Another child, F14 (11110), used multiplication to capture different relationships; he was the only child who answered all problems correctly in AL.

We now examine the responses drawn by F17 (10000) for P2 and P4. In BL, F17 tried to add 30 18 times for P2. To solve P4, he started to accumulate 8s and 60s in parallel but ended up in the middle. He was struggling with repeated additions in BL. On the contrary, F17 began to use multiplication in AL (Fig. 3).

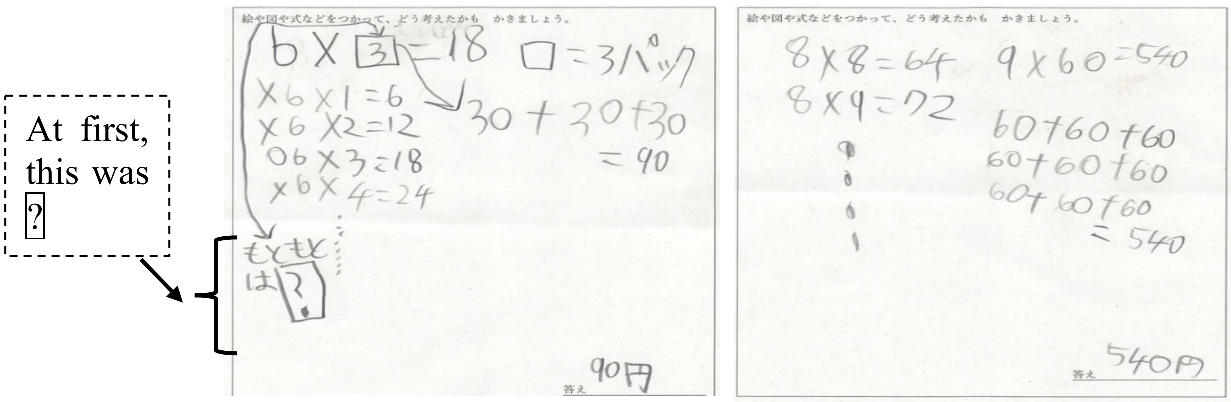



Fig. 3: F17's drawn responses to P2 (left) and P4 (right) in AL

To solve P2, F17 explored the relationship between 6 and 18 eggs. He attempted a multiplication table of 6 and found that □ was 3, and wrote "□ = 3 packs." However, he used addition to obtain an answer for the price. A similar attempt can be seen for P4. The four dots indicate that he was also thinking about 8 ×1, 8 ×2, and so on. Then, starting from 8 × 8, he derived "×9." This time, to determine the price, he used the multiplication "9 × 60" besides addition. As such, his responses indicate progress in his thinking strategy from BL to AL, and from P2 to P4 in AL.

## DISCUSSION AND CONCLUSION

By learning multiplication, most of the children were able to solve P1. However, the children continued to have difficulty with P2–P5. Because these are many-to-many correspondence problems, our results agree with Vanluydt et al. (2020). To understand the children's change in more detail, we examined the uses of multiplication they had learned and found that there was diversity in how children who had previously been adding, subtracting, or counting began to use multiplication. Multiplication learned by the children does not differ significantly at the textbook level. Nevertheless, how the children used it in solving the problem reflected the meaning they gave to multiplication. In addition, a child's use of multiplication reflected the quality of proportional reasoning that the child had previously developed.

The codes "form" and "simplification" capture some children's meaning of multiplication. They provided a framework for identifying the multiplicative aspect of a problem for children who had previously had difficulty understanding the problem or focused only on the number of eggs. In particular, "simplification" may be one of the earliest responses using multiplication to solve many-to-many correspondence problems. Here, the children failed to recognize a composed unit but simplified it as a unit ratio, changing the problem to a one-to-many correspondence. This can be a unique usage in the process of expanding the view from one to two quantities.

The code "relation" is seen in children with various BL patterns. It is worth scrutinizing further because our preliminary analysis indicates the difficulty of building a composed unit in a multiplicative manner. The use of multiplication to determine the relationship between the number of eggs was observed in a relatively large number of children. Meanwhile, when the relationship between the price in the problem statement and the price to be found is captured as multiplication, the composed unit needs to be considered. For example, in P4, when 60 ×9 is concerned, "9" is derived from the relationship between the number of eggs (8 × 9). In other words, the composed unit (8 eggs: 60 yen) is attended to, and the two quantities are both multiplied by 9. Few children used multiplication to find price as well; in particular, none of the children belonging to the 00000 pattern in BL did so. In this connection, by looking over all the drawn responses of two children who were able to build up in a multiplicative manner and who were not for P4, we notice two differences: first, how abstractly they expressed the number of eggs shown in the problem text (by a number or by drawing individual circles), and second, how clearly they expressed the correspondence between the number of eggs and the price to be found. Further analysis of the progression from additive to multiplicative treatment of composed units is required.

Even in discrete quantity problems, which may be relatively easy as a missing-value type, our results show that children's progression of EPR is complex. Moreover, it is closely connected with the meaning they develop for the learned concept. The implications of the results for teaching multiplication should be further explored.

## ACKNOWLEDGMENT

This project was supported by JSPS: KAKENHI (20H01671, 21K02593, 22K02660).

## REFERENCES

Harel, G., & Confrey, J. (Eds.). (1994). *The development of multiplicative reasoning in the learning of mathematics*. Albany, NY: State University of New York Press.

Hino, K., Kato, H., & Ichikawa, H. (2022). Proportional reasoning of lower grade students through learning trajectory and lesson study. In C. Fernandez, S. Llinares, A. Gutierrez, & N. Planas (Eds.). *Proc. 45th Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 4, p. 230). PME.

Hino, K., Ueno, Y., Tajima, T., Sakurai, A., & Ichikawa, H. (2023). A study of proportional reasoning in second-grade students before and after learning multiplication. *Proceedings of the 56th fall annual meeting of the Japan Society of Mathematical Education* (pp. 353-356). (in Japanese).

Isoda, M., & Olfos, R. (Eds.) (2020). *Teaching multiplication with lesson study: Japanese and Ibero-American theories for international mathematics education*. Springer.

Kaput, J. J., & West, M. M. (1994). Missing-value proportional reasoning problems: Factors affecting informal reasoning patterns. In G. Harel & J. Confrey (Eds.), *The development of multiplicative reasoning in the learning of mathematics* (pp. 237–287). Albany: State University of New York Press.

Lamon, S. J. (2005). *Teaching fractions and ratios for understanding* (2nd ed). Mahwah: Lawrence Erlbaum Associates.

Lobato, J., Ellis, A. B., Charles, R. I., & Zbiek, R. M. (2010). *Developing essential understanding of ratios, proportions, and proportional reasoning for teaching mathematics in grades 6–8*. Reston: National Council of Teachers of Mathematics.

MEXT (2012). *Summary of expected future efforts based on the results of the National assessment of academic ability over the past four years*. (in Japanese).

MEXT (2018). *The course of study for elementary school mathematics*. Tokyo, Japan: Nihon Bunkyo Shuppan. (in Japanese).

Singh, P. (2000). Understanding the concepts of proportion and ratio constructed by two grade six students. *Educational Studies in Mathematics, 43*, 271–292.

Vanluydt, E., Degrande, T., Verschaffel, L., & Van Dooren, W. (2020). Early stages of proportional reasoning: A cross-sectional study with 5- to 9-year-olds. *European Journal of Psychology of Education, 35*, 529–547.

# ‘DODGING THE BULLET’: CONSTRAINTS ON THE USE OF DERIVATIVES IN MECHANICS COURSES.

Mathilde Hitier and Alejandro S. González-Martín

Université de Montréal

*The calculus notion of derivative plays a central role in kinematics. However, previous research shows that at the college level, instructors rely more on ready-to-use formulas than on covariational reasoning when teaching kinematics. In this paper, we identify constraints placed on mechanics teachers when working with the derivative in a kinematics context. Our results indicate that the traditional separation of knowledge into different branches (e.g., mechanics and differential calculus) has a strong impact on the teaching practices of mechanics instructors. Specifically, what students learn (or do not learn) in their calculus courses places limitations on mechanics teachers, restricting their use of tools from calculus to fully develop students’ understanding of motion and instantaneous rate of change.*

## INTRODUCTION

Stoffels et al. (2022), among others, have pointed out that “physics is the prime example of mathematical applications and the stimulus for mathematical theories”. More specifically, the authors identify dynamics as “the most common content that connects differential calculus and physics, especially in undergraduate mathematics and physics”, mainly based on the fact that “the concept of instantaneous velocity [is] a standard example of the instantaneous rate of change” (p. 235). They note that “differential calculus in physics is usually introduced in the context of kinematics” (p. 237). Their recent literature review of physics and mathematics education research highlights the important role that calculus plays in the study of kinematics and more advanced physics concepts. They also note that some students experience difficulties with calculus in both a mathematics and physics context. For instance, in mathematics, students have difficulties with the derivative as slope of the tangent line linked with the limiting process involved, and most students fail to fully grasp how to apply differential calculus in a physics context.

Our previous work shows how instructors’ teaching practices differ when addressing the derivative in a calculus context versus a kinematics context. In calculus classes, algebraic expressions are used to arrive at differentiation formulas. When it comes to kinematics, on the other hand, teachers tend to employ ready-to-use formulas, sidestepping notions of covariation and instant rate of change (Hitier & González-Martín, 2022a). These differences in teaching practices have an impact on students, who struggle to see the connection between the different processes. The result is a compartmentalisation of techniques: students rarely use procedures learned in calculus to solve kinematics problems, and vice versa (e.g., Hitier & González-Martín, 2022b). Research concerning the use of the derivative in physics courses (and mechanics



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 41-48). PME.

courses in particular) is growing, but is still scarce, with a lack of studies focusing on the teachers' perspective.

In recent years, the practices of postsecondary teachers have been getting more attention from researchers (Winsløw et al., 2018). While some have examined the differences in teaching practices by looking at different instructors of the same course (e.g., Wagner & Keene, 2014), there is a lack of research examining the impact of external constraints on postsecondary teachers' practices (González-Martín, 2018). Working with school teachers, Robert and Rogalski (2002) suggested that some variations in the teachers' practices depend on factors outside of their personality and training, positing that external constraints can make regularities appear among teachers when they adapt their practices to these constraints. This paper aims at better understanding this phenomenon as it applies to mechanics teachers at the college level.

## THEORETICAL FRAMEWORK AND RESEARCH QUESTION

In our research, we take the point of view of the Anthropological Theory of the Didactic (ATD), which sees all human activity as institutionally situated. Within this framework, we consider mechanics and calculus as different institutions. A key analysis tool of ATD is that of praxeology, which is made up of two interrelated blocks. The practical block consists of a type of task and a technique to perform this task. The practical block is supported by the theoretical block, which includes a rationale (called a technology) that justifies the technique and is embedded in a larger theory (e.g., Chevallard et al., 2022). ATD identifies various *levels of didactic codeterminacy* that allow us to pinpoint different constraints affecting the organisation, teaching and learning of content (see Figure 1 for the full scale). Each level imposes a set of constraints on and supports for the didactic system. The lower levels include *questions* (the study of specific types of tasks, also called *subjects*), *themes* (different tasks connected by the same rationales), *sectors* (collections of several themes), *domains* (groups of themes), and *disciplines*. In our study, we observe that decisions at the *schools* level affect the order in which courses are taught. We see mathematics as a discipline and differential calculus as a domain, with the study of the derivative being a sector. Physics is a separate discipline, with mechanics being a domain and kinematics a sector.

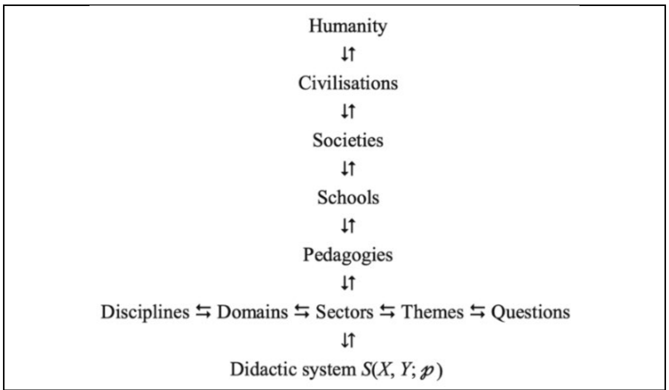



Figure 1: Scale of didactic codeterminacy (Chevallard et al., 2022, p. 174).

In Hitier & González-Martín (2022a) we compared the use of the derivative in calculus and mechanics courses, particularly at the levels of questions, themes, and sectors. This allowed us to identify inconsistencies between teaching practices in the two courses related to the use of the derivative. Our study revealed that even though velocity and acceleration are defined as derivatives both in calculus and mechanics, the mechanics course employed rationales primarily centred around the use of kinematics equations, neglecting conceptual aspects (Hitier & González-Martín, 2022a). We are interested in studying mechanics teachers' views on this practice and the reasons behind it. In this paper, we investigate the constraints imposed on mechanics teachers that arise from practices in calculus classes. We consider these to be external constraints, since mechanics teachers do not have any control over them. We also seek to situate these constraints at the various levels on the scale of didactic codeterminacy.

## METHODS

This paper is part of a larger research project at a large Canadian college (Cégep) in Quebec. Cégeps are postsecondary institutions that Quebec students must attend prior to pursuing university studies. In this Cégep, science students usually take differential calculus (Cal I) and mechanics in their first term, followed by integral calculus (Cal II) in their second term. They also have the possibility of taking an elective multivariable calculus course (Cal III) in their fourth and final term.

Our research relies on three different sources of data: calculus and mechanics textbooks (Hitier & González-Martín, 2022a), the students themselves via online questionnaires and task-based interviews (e.g., Hitier & González-Martín, 2022b), and calculus and mechanics teachers through interviews and classroom observations. In this paper, we draw on our interviews with four mechanics teachers who volunteered to participate in the study (see Figure 2).

| Teacher | Teaching experience at Cégep level | Frequency of teaching the course |
|---|---|---|
| P1 | 10 to 15 years | Almost every fall term |
| P2 | 15 years | Almost every fall term |
| P3 | 15 years | Almost every term |
| P4 | 20 years | At least once a year |

Figure 2: Background information on the mechanics teachers.

The semi-structured interviews took place online at the end of both the Fall 2020 term (P1 to P3) and Winter 2021 term (P4). The interviews lasted from 40 to 100 minutes and their recordings were transcribed prior to our analysis. Both authors coded independently and any minor differences were resolved. We then analysed the transcripts to extract the mechanics teachers' justifications for their praxeologies. For this paper, we focus on those that are influenced by the teacher's perceptions of the content (or lack thereof) taught in the Cégep's mathematics courses, which is seen as imposing constraints on their teaching practices. These constraints were then classified according to their level of codeterminacy. In the next section, we start by describing constraints originating at the higher levels (beginning with the school level) continuing down the scale to the sector level. This allows us to determine how the content taught in calculus courses affects the teaching practices in the mechanics course.

## ANALYSIS

All four participants identify constraints at the **school level**; namely, the way the Cégep organises its course schedule and separates mathematics and physics into two disciplines, studied concurrently. For instance:

> I would love to be able to talk about physics, like talking about derivatives and integrals. But […] when the students come into the classroom, I can't assume that they have that math background. (P3)

P1 states his belief that "being comfortable with the tools of math and then doing physics […] is better". Furthermore, the separation of these disciplines, as well as the Cégep's scheduling of the courses, result in what P2 calls a "misalignment" of the calculus and mechanics courses:

> So that was the biggest problem for us: it was that misalignment, the fact that we were using tools [from calculus] they hadn't developed yet. So the trick, and it was a trick, as a sick trick, but we had to somehow defer using those tools even though they were being talked about. (P2)

P2 specifies that "[teachers] always had to defer talking about utilising the derivative until it appeared in [the students' calculus] courses, usually around like week five or later, even". Therefore, the mechanics teachers who are concerned about using derivatives before they are presented in the calculus class, feel forced to organise their content accordingly. The participants give other examples of how this misalignment influences their teaching practices, with P4 explaining how it limits his ability to explore material in greater depth:

> P4: Maybe, if I had time, I would […] provide them with some prerequisites. But not like a math teacher, like a physics teacher. Which means […] using physics examples, not math. […] I would go for that at the beginning. It would help, I think.
>
> […]
>
> Researcher: If you had […] time [to take care of this] introduction yourself […] would […] you […] do something different with respect to the use or interpretation of the derivative later on?
>
> P4: […] Of course! I would, because […] the concept of derivative [beyond formulas] would help me to cover more material and deeper material.

In addition to complaining about the order in which calculus and mechanics courses are taught at their Cégep, the participants also note the impact of spreading content concerning derivatives over two different courses.

The **discipline level** sheds light on the epistemology of the disciplines and the participants' views on their relationship, which resonates with Redish and Kuo (2015). Participants complain that the content taught in the mathematics courses in general, and calculus in particular, does not adequately prepare students for what they encounter in their physics course. For instance, P1 states that "numbers need units to have any sort of meaning", and that this creates difficulties for students in "apply[ing] [their math knowledge] when they write physics equations, because they're used to doing math with numbers and not numbers that have units". Like the participants in Redish and Kuo (2015), our teachers also use the metaphor of language. P1 considers "math [to be] the language […] that physics is written in", and adds that "[students] do not have enough of it", which P2 sees as problematic:

> The problem we have as physicists is that we want to teach them how to think physically, but the language is mathematics, so we're […] stuck in the job of teaching two courses: here's how you use mathematics, how you write mathematics, and this is how you do mathematics, and now we're going to use that to talk about this thing you don't understand.

The separation of the disciplines at the school level also impacts the organisation of content at the **domain level**, in particular mechanics content:

> I can't invest any more time into trying to develop the mathematical machinery. I just have to assume that that idea is there, and it's been developed […] in their calculus course, because then […] I have to run on and […] talk about the physical concepts, right. So I can't go back and revisit defining the derivative […] That was the weakness of the misalignment in the two courses. […] I was using the derivative and expecting them using the derivative in ways they hadn't been taught in calculus yet, so there was a lot of hand waving and hoping that "oh, you know, eventually you will get this in your cal course" but uh we didn't rely upon it too much but I still had to pretend it was there so… I was sad. (P2)

The participants believe the separation of disciplines and the misalignment of courses influence many of their teaching practices: they opt to quickly present the mathematics underlying a mechanics notion, focusing on ready-to-use formulas that will help students perform calculations. With this reliance on formulas, students miss out on the deeper mathematical meaning behind the physics notions under study, which the teachers see as a consequence of students' unfamiliarity with derivatives in calculus.

The use of vectors at various levels—in physics in general (at the discipline level), and in kinematics in particular (at the **sector level**)—reveals inconsistencies beyond those identified in our textbook praxeological analysis (Hitier & González-Martín, 2022a). More specifically, although Cal I is a single-variable differential calculus course, "you must understand vectors to talk about motion" (P2):

> Velocity is a vector, and a vector always has a magnitude and a direction […] and […] there's no scalar equivalent of acceleration, right […], there's always a direction for acceleration. (P3)

P2 explains why this is a major constraint on using the derivative in the mechanics course:

> We're […] talking about 1D motion, but […] immediately, we're talking about vectors […] we're differentiating vectors … vectors that are functions in time … so […] in a certain sense [students] never get that background unless they take Cal III and even then […] that's the fourth term. We're in term one so … […] we're screwed!"

The Cégep's separation of mathematics and physics has another consequence at the sector level: “[Students] can do the derivatives, so [teachers] can give them a position function and we can take the derivative, but then we can only talk about … going the other way [instead of using integration to retrieve a position function]” (P1). In short: “We’re not doing integrals, but we can do derivatives” (P4). Tasks in mechanics include not only finding the velocity or acceleration functions through differentiation, but also finding the velocity function from the acceleration function (Hitier & González-Martín, 2022a). Therefore, the sector of 1D kinematics in physics is linked to the domain of multi-variable calculus, as well as to two other mathematics domains: differential and integral calculus. The constraints stemming from this situation might explain why “the models [used] in the class can only do […] constant acceleration” (P3), as it allows students to retrieve the velocity and position geometrically and algebraically, without involving the use of formal integration. While professional physicists may engage in these activities and techniques, students cannot develop them in their mechanics course for two reasons: students are introduced to differential and integral calculus (two domains within the same discipline) in two separate courses, and the order in which courses are taught at the Cégep (school level) is not ideal. P2 explains how this affects the way he addresses the relationship between the different motion functions:

> So I get them to do it algebraically, functionally, even though they haven’t seen calculus yet, and then graphically and then that’s coupled with doing the motion diagrams… […] We couple this with activities in the lab where they’re actually watching and measuring motion and one hopes that through all this activity and studying that eventually the concept of motion starts to make sense to them.

In particular, P2 uses motion diagrams to bypass the limiting process by working discretely:

> This notion of what we call motion diagrams, where you take a continuous motion and you break it into discrete chunks, and you take a look at “well over this interval there’s a change, over this next interval there’s another change”. And by breaking it up into these discrete chunks, you didn’t have to worry about […] “here’s a continuous function, I have to talk about this rate of change, I have to introduce limits blah blah blah”, all that stuff. So we’re able to sort of, you know, dodge the bullet there. It’s like: “I’m gonna teach you the derivatives, but not really! Here is some discrete math, yeah!” Okay, so we draw our dots […], even though it’s disingenuous in the sense that we’re not teaching them how to think about velocity properly…”

Only half of our teacher participants take time to really work with motion diagrams, but they all put a particular emphasis on graphs in order to avoid relying on the calculus:

> Unfortunately, we start teaching this when they did not do derivatives at all, at all, at all […] so […] that’s why my emphasis was […] based on considering graphs […] because if you don’t understand conceptual, […] like abstract, so better consider visual […] slopes […] for derivatives and areas for integrals […] Moreover as they do not understand the concept of derivatives, […] … yet. Which means the derivative to them it’s […] something abstract, not like a local characteristic and time rate, etc., etc. (P4).

In short,

> the graph is math but it is easily digestible math […]. There's no numbers, so students are not as afraid of it. So I guess […] I get my math, but without having to make sure that they actually have the background for the math (P3).

The relationship between the sector of 1D kinematics in physics and the two domains (differential and integral calculus) of mathematics strongly impacts the mechanics teachers' practices, forcing instructors to use graphs or to take a discrete approach to the study of motion. This, despite the fact that these approaches fail to foster an optimal way of envisioning motion, according to the teachers. As a result, they feel forced to limit the study of motion, focusing on simple cases that are studied without explicitly using derivatives. As a consequence, these strategies (which more or less bypass the derivative) lead P4 to state that 1D kinematics "[is] more or less simple […] it's a linear motion with constant acceleration, so it's about using four equations of kinematics".

## DISCUSSION AND CONCLUSION

In this paper we investigate how separating content related to derivatives into two different courses (calculus and mechanics) impacts the study of motion in mechanics courses. We have not found other papers in the mathematics education literature addressing these issues; therefore, we believe our study can shed light on how content taught (or not taught) in calculus courses can affect teaching practices in other courses.

Our data reveal a number of constraints that clearly influence the teaching practices of mechanics instructors, and all our participants agree that what they do in their courses does not reflect common practices in physics. The timing of when students begin learning about derivatives in their calculus classes means that mechanics teachers cannot introduce mechanics content based on covariation. Aside from quickly defining derivatives before deriving the kinematics formulas, instructors mostly resort to ready-to-use formulas that do not allow students to fully grasp the concept of movement. Furthermore, the teachers cannot use integrals explicitly, which means they can only tackle simple problems (constant acceleration) that allow for the use of basic graphs ("easily digestible math"). Moreover, the fact that students have only been exposed to one-variable calculus prevents the instructors from teaching the study of movement from a vector perspective, and they feel that students "never get that background" to adequately study motion. The teachers also consider practices in calculus (and mathematics in general) related to plain numbers as creating difficulties in addressing magnitudes and units, which is also an issue mentioned by Stoffels et al. (2022).

Our results seem to agree with the work of Robert and Rogaslki (2002) at the school level. We see some regularities in our participants' adaptations to their teaching practices as a result of external constraints. The location of these constraints on the scale of codeterminacy reveal that the mechanics teachers have little to no control over them and are forced to adjust their teaching practices as a result, leading to frustration. We plan to pursue this study further in order to better understand teaching practices in mechanics courses, and to gather information that can help calculus teachers improve their practices. Our study could also lead to recommendations regarding the sequence of calculus and mechanics courses at the college/Cégep level. We also believe that this type of study will shed light on practices in other courses/disciplines; the possible impacts of this research and the identification of constraints may lead to solutions that will benefit students and improve learning outcomes in other fields of study. This will also benefit teachers, who may feel less frustrated if their course content more accurately reflects professional practices.

**References**

Chevallard, Y., Barquero, B., Bosch, M., Florensa, I., Gascón, J., Nicolás, P., & Ruiz-Munzon, N. (Eds.). (2022). *Advances in the Anthropological Theory of the Didactic*. Birkhäuser Cham. https://doi.org/10.1007/978-3-030-76791-4

González-Martín, A. S. (2018). Les cours préparatoires à l'université, au Québec : the same, but different? In M. Abboud (Ed.), *Actes du colloque Espace Mathématique Francophone 2018 (EMF2018)* (pp. 548–556). Université de Cergy-Pontoise.

Hitier, M., & González-Martín, A. S. (2022a). Derivatives and the Study of Motion at the Intersection of Calculus and Mechanics: a Praxeological Analysis of Practices at the College Level. *International Journal of Research in Undergraduate Mathematics Education*, *8*(2), 293–317. https://doi.org/jq4x

Hitier, M., & González-Martín, A. S. (2022b). "It all depends on the sign of the derivative": A praxeological analysis of the use of the derivative in similar tasks in mathematics and mechanics. In J. Hodgen, E. Geraniou, G. Bolondi, & F. Ferretti (Eds.), *Proceedings of the Twelfth Congress of European Research Society in Mathematics Education (CERME12)* (pp. 2421–2428). ERME / Free University of Bozen-Bolzano. https://hal.archives-ouvertes.fr/hal-03750626

Redish, E. F., & Kuo, E. (2015). Language of Physics, Language of Math: Disciplinary Culture and Dynamic Epistemology. *Science & Education*, *24*(5), 561–590. https://doi.org/f7fgrn

Robert, A., & Rogalski, J. (2002). Le système complexe et cohérent des pratiques des enseignants de mathématiques : Une double approche. *Canadian Journal of Science, Mathematics and Technology Education*, *2*(4), 505–528. https://doi.org/dp5fwr

Stoffels, G., Witzke, I., & Holten, K. (2022). Comparison: Differential Calculus Through Applications. In F. Dilling & S. F. Kraus (Eds.), *Comparison of Mathematics and Physics Education II: Examples of Interdisciplinary Teaching at School* (pp. 227–242). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-36415-1_17

Wagner, J. F., & Keene, K. (2014). Exploring differences in teaching practice when two mathematics instructors enact the same lesson. In T. Fukawa-Connelly, G. Karakok, K. Keene, & M. Zandieh (Eds.), *Proceedings of the 17th Research in Undergraduate Mathematics Annual Conference* (pp. 322–335). MAA

Winsløw, C., Gueudet, G., Hochmuth, R., & Nardi, E. (2018). Research on university mathematics education. In T. Dreyfus, M. Artigue, D. Potari, S. Prediger, & K. Ruthven (Eds.), *Developing Research in Mathematics Education: Twenty years of communication, cooperation and collaboration in Europe* (pp. 60–74). Routledge.

# CO-CONSTRUCTING AN IMAGE OF VALUED MATHEMATICS TEACHING: NOTICING AND NAMING STRENGTHS IN VIDEO RECORDS OF PRACTICE

Anna Hoffmann, Jen Munson

Northwestern University

*While much of the literature on supporting teacher noticing in video records of practice advocates for a neutral approach, recent research on classroom-based noticing has pointed to the value of applying a strength-based lens. In this study, situated in a video-based professional development program in which teachers were asked to attend to strengths, we explored what teachers identified as strengths in video records of mathematics teaching and to whom they attributed these strengths. Analysis of six discussions identified five themes: (1) Designing and engaging in the mathematical space, (2) Designing and engaging in the discursive and collaborative space, (3) Establishing norms, (4) Growth, and (5) Engagement. This study suggests affordances for designing teacher noticing protocols for video records with a strengths-based lens.*

## FRAMING

What professionals notice and highlight in artifacts and records that represent their practice reflects their professional vision, their goals, and values as professionals (Goodwin, 1994). Recent research has explored how learning to notice student thinking while assigning value to it and acting upon identified strengths represents an equitable approach to mathematics teaching (e.g., Jilk, 2016; Louie, 2018). That is, by noticing student thinking as an asset, teachers position students as competent and capable learners with understanding that can be built upon. However, research has yet to explore the outcomes of strengths-based noticing in mathematics teacher development programs when the lens for looking is open not just to student thinking but all that might be captured in videos of mathematics teaching and learning in the classrooms.

A robust body of research indicates that mathematics teachers' professional development programs that focus on noticing and discussing video-taped classroom events contribute to mathematics teacher knowledge and practice (e.g., Gaudin & Chaliès, 2015; Santagata et al., 2021; van Es & Sherin, 2008). Researchers typically advocate neutral, non-judgmental discussion of classroom events as it enables slowing down and deeply analyzing and considering practice (e.g., Coles, 2019; Karsenty et al., 2019; van Es, 2011). Hence, there is a potential tension between strengths-based and neutral noticing of video-taped classroom events, a tension that requires teacher educators' attention while designing and facilitating mathematics teacher development programs.

We argue that what teachers notice as strengths in video records of practice partially reflects what they see as features of effective mathematics instruction. Thus, noticing, naming, and discussing strengths in video records of practice represents one way that



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 49-56). PME.

teachers' professional vision becomes visible and open to negotiation. Negotiating not just what is seen, but what is valued in mathematics teaching and learning is often a central goal of professional learning. As a first step to studying the affordances of noticing strengths in video records of mathematics classrooms for teacher development, we explore the following research question: *What do teachers identify as strengths in video records of mathematics teaching and to whom do they attribute these strengths?*

## METHODS

### Study design setting, and participants

Our research took place in a two-year professional development program in the US for practicing secondary mathematics teachers at the beginning of their careers (2-7 years of experience) that was designed to promote equitable mathematics instruction using a collaborative, inquiry-based model (Cohen et al., 2014). The program included a two-week residential institute each summer, followed by both individual and small-group virtual coaching sessions throughout the school year. Our current study focuses on the small group sessions, which provided an opportunity for teachers to see into one another's classrooms through video records of their lessons and discuss emergent issues regarding their practice.

Twenty-two teachers participated in the cohort under investigation, divided into seven small groups of 3-4 teachers plus a facilitator. Each group met three times per year. In each small group session, each participating teacher contributed a video clip from their classroom for group discussion, which the other group members viewed in advance and tagged with comments and questions. The discussions of each video, lasting approximately 20-25 minutes, followed a conversational protocol (McDonald et al., 2007). First, the video-sharing teacher described the context of the video clip and addressed any contextual questions posed by the other participants (approximately 5 minutes). Second, the facilitator invited each person, beginning with the video-sharer, to name "strengths and highlights" from the video (approximately 5-10 minutes). Third, the video-sharer was prompted by the coach to name a question emerging from the video clip, which the group then discussed for the remaining time (approximately 10-15 minutes). The design differed from what is typical in discussions of video records of practice as teachers were invited to name strengths and highlights of the videos; that is, their noticing was not neutral, but rather intended to locate features of the video that they deemed valuable. During the discussion, participants referred to strengths and highlights interchangeably; as such, we do the same.

### Data selection and analysis

Over the two years of the program, 41 small group sessions were recorded on the video conferencing platform used by the program, resulting in 127 discussions of videos. Each included noticing and naming strengths of the video being discussed. Each small group session was professionally transcribed. In addition to the video records and the transcripts, the facilitator took notes in the form of a structured memo which followed the conversational protocol and was provided live to participating teachers for reference. Upon repeated viewing of the videos and concurrent reading of the transcripts and facilitator memos, the research team determined that the facilitator memos were in strong alignment with the transcripts, often including verbatim descriptions of participants' contributions to the discussion. We opted to analyze the memos, using the videos and transcripts to clarify any ambiguities. In this preliminary analysis designed to develop an inductive framework for what teachers noticed as strengths, we analyzed data from six small group meetings, where a total of 18 videos were discussed. The sessions were chosen from the beginning and ending of the program to reflect possible changes over time in the focus of the strengths and were selected to include a wide number of participants (16 teachers) to reflect possible differences among participants.

To analyze what teachers noticed as strengths, we created a spreadsheet that listed each instance of a named strength that appeared in the memos. We open-coded the strength by describing its main idea, and noted the role of the participant who named the strength (i.e., video-sharer, video-viewer, or facilitator) and to whom the strength was attributed (i.e., teacher, individual student, group of students, whole class, or task). When an instance addressed multiple strengths, we further segmented that instance into individual ideas and open coded each separately. We continued open-coding iteratively and comparatively, creating additional codes where needed, refining existing codes, and collapsing related ones. The first author coded all the data and developed an initial coding scheme (i.e., code names and definitions). Based on this, the second author coded 10% of the data. Then we discussed disagreements until consensus was reached on a coding scheme that included 17 codes. Finally, we grouped the codes into major themes, described in the findings.

## FINDINGS

Our analysis revealed 133 instances of naming strengths in the small-group video discussions that were analyzed. Fifty-one of these instances were named by the video-sharing teacher, 57 instances were named by video-viewing teachers, and 24 instances were named by the facilitator.

The participants assigned the strengths that they noticed to different entities, which included but moved beyond noticing strengths in students' contributions. As Table 1 indicates, nearly half (n=61 of 133) of the strengths were assigned to the teacher who appeared in the video, the video-sharing teacher. About a third of those (n=19) were self-assigned by the teacher who shared the video. More than one-fourth (n=36) of the strengths were assigned to students, most were assigned to a group of students (n=24), and the remainder were assigned to a particular student (n=12). Nearly one-fourth of strengths were assigned to the class as a whole, meaning the teacher and students together. A small number of the strengths (n=7) were assigned to the mathematical task that was discussed in the video.

| Entity to whom strength was assigned | Count | Percent |
|---|---|---|

| | | |
|---|---|---|
| Teacher (by video viewer or coach) | 42 | 32% |
| Teacher (self-assigned by video sharer) | 19 | 14% |
| Class (students and teacher) | 29 | 23% |
| Group of students | 24 | 17% |
| Individual student | 12 | 9% |
| Task | 7 | 5% |
| Total | 133 | 100% |

Table 1: Number and percentage of strengths assigned to different entities.

We grouped the strength codes into five themes: (1) Designing and engaging in the mathematical space, (2) Designing and engaging in the discursive and collaborative space, (3) Establishing norms, (4) Growth, and (5) Engagement. The themes are not distinct but connected and overlapping. Yet, they highlight the different foci of the strengths that the participants named. Each theme could apply to multiple actors. For example, the code of *mathematical persistence* which was grouped under the theme of *designing and engaging in the mathematical space* was assigned sometimes to students, recognizing their persistence despite their struggle, other times to teachers, recognizing their scaffolding students' mathematical persistence, and occasionally to a task, recognizing a feature of the activity that encourages mathematical persistence. In the following, we describe and exemplify these themes.

**Designing and engaging in the mathematical space**

Within the theme of *designing and engaging in the mathematical space,* teachers noticed as strengths the ways that both teachers and students engaged in mathematics. Participants attended to the ways that the teacher designed and facilitated the mathematics by selecting worthy tasks; providing productive prompts, questions, and explanations; pressing on valued mathematical practices such as justification and perseverance; and thoughtfully and flexibly responding and scaffolding students in their mathematical work. For example, a strength that addressed the design of the mathematical space stated: "The task encouraged flexible thinking and open-mindedness because there was no clear best proof." An example of pressing on valued mathematical practices such as explaining why things are true and facing struggles while doing so stated: "Megan really held their feet to the fire around the 'why'. Students were getting frustrated, but not giving up."

Teachers also noticed the assets that students brought to mathematical work, including their knowledge of mathematical concepts and methods, their engagement in mathematical problem-solving and sensemaking, and the way they expressed these in front of the teacher and their peers. One participant stated, "It was clear that they have the basic understanding and were analyzing it to another level. They addressed higher-level areas in their feedback," drawing attention to students' mathematical knowledge.

Another participant appreciated clearly communicating a mathematical idea: “The student who was able to articulate that they know it works because of these four examples, but don’t know how to prove it for all functions.”

**Designing and engaging in the discursive and collaborative space**

The *designing and engaging in the discursive and collaborative space*, included teacher noticing of assets that students and teachers brought to classroom interactions. It included the ways that teachers facilitated discussion among small groups of students and the whole class and how the teacher talked with individual students. For example, one participant described “Alex did a good job of not giving his opinion or letting students know what his opinion is,” while another stated, “Megan always gives a question to leave the group with before she leaves.”

Teachers paid specific attention to the ways that students collaborated to solve problems, shared their ideas, listened to one another, and built on one another’s thinking. For example, one participant described, “Students were really willing to listen to each other,” while another stated, “When Kevin started talking, students physically turned to the front with attention. It makes a big difference for a student to be at the center of attention.” Teachers also noted the ways that teachers used discourse to elevate student ideas. For instance, a discussant noted that the video-sharing teacher “did a good job about giving public credit to student ideas.”

**Establishing norms**

The theme of *establishing norms* included strengths which focused on the establishment or existence of valued norms of the mathematics classroom. Participants focused on evidence that students knew what was expected of them within activity, that students trusted one another or felt comfortable to contribute, that students treated mistakes as normative, or that there was a shared understanding of what it meant to do mathematics. For instance, one teacher described, “The classroom culture had such collective ownership over the problems. There was such a lot of safety.” One video sharer described the following strength related to norms: “I’m working toward helping students be more aware of their behavior. I am trying to take the stance that students are still learning executive functioning. We have to show them other areas where they can redirect their energy.” Additionally, we included in this theme strengths of personal characteristics that support valued norms, for instance, “I was so impressed by how calm and focused you were on what you wanted at all times.”

**Growth**

The theme of *growth* included strengths that focused on a comparison between past and current events from the perspective of improvement over time. The growth teachers noticed included increased student participation (either individually or collectively), evidence of the teacher’s prior efforts to implement routines or practices, and shifts in teachers’ use of practices, particularly those discussed in the professional development program. For instance, one participant noted a shift in eliciting student thinking when describing, “There is more eliciting in this video than in previous videos. Really seeing kids explain their ideas.”

Most comments that appeared in this theme were made by the video-sharing teachers, as they were the most familiar with the history of the recorded class. For instance, one video sharer stated in reference to a student, “It was the first time he stayed seated for the whole class.” Yet, several comments were made by the video-viewing teachers and the facilitator based on previous videos that the group watched, such as the earlier comment about increased eliciting.

**Engagement**

The theme of *engagement* highlighted moments when students actively, promptly, or widely engaged in the mathematical activity in the class. For example, one participant stated in a typical comment in this theme, “All of the students were on task the whole time.” Strengths of this theme typically appeared at the beginning of a teacher’s talking turn, using it as an opportunity to name other strengths with diverse foci that explain what assets of mathematics teaching and learning enabled the engagement. For instance, the next citation from a discussion transcript exemplifies a video-viewing teacher who first noted the ‘energy’ in the class and then went on to attribute students’ engagement to the nature of the question the teacher had posed, “I was going to say that the energy… these kids, their butts are out of their seat waving their arms… I guess for me I didn't think about it as a question that would ever spark so much.” In doing so, this teacher connected student engagement to the teachers’ design of the mathematical space.

## DISCUSSION

To investigate the affordances of noticing and naming strengths in video records of mathematics instruction for mathematics teacher development, we explored what teachers identify as strengths in their own and their peers' videotaped mathematics lessons and to whom they attribute these strengths. We acknowledge that the image captured by the strengths that teachers noticed in our research is partial (e.g., not every strength is noticed, the videos are limited in the scope of the strengths that they present) and is shaped by the focus of the given professional development program. Yet, the strengths that participants identified paint a collective picture of their professional vision (Goodwin, 1994) of mathematics teaching, which includes a wide range of activities and actors in the mathematics classroom. Participants valued as assets designing for and engaging in rich mathematics through discourse, collaboration, and the establishment of norms consistent with these goals. While noticing high levels of engagement was prominent in the data, we found that participants did not attend to engagement as an end to itself but rather an entry point to identifying contributing factors and other assets. Finally, the participating teachers were able to notice in one another, their students, and themselves evidence of incremental growth toward valued goals.

Collectively, our findings suggest that naming strengths as part of a video-viewing protocol supported the group to make visible a professional vision that explicates valued features of mathematics teaching and how these features look in complex and diverse classrooms settings. We do not claim that strengths-based or neutral noticing is always preferable; rather, we argue that this design choice involves careful consideration by teacher educators and professional development designers alongside other design decisions. For instance, we posit that use of a strengths-based lens may be influenced by whether the video being viewed and discussed was drawn from a participating teachers' classroom, as it was in our data. In studies in which the videos have been selected by facilitators to show unknown teachers and students with the deliberate aim of provoking inquiry and analysis, Schwarts and colleagues (2022) argued that the videos must be framed to avoid implications that they represent exemplary practice. They argue that perceiving artifacts as models to be imitated could prevent deep analysis of the practices they represent. In our study, where the teachers each contributed videos with the aim of mutual learning and growth, the conversations did not succumb to this pitfall. Attending to and naming strengths functioned not to lionize the teacher, but to sharpen the lens for noticing teachers' and students' earnest efforts and growth.

In future analyses, we aim to expand our examination to the small group discussions as a whole and explore the ways that noticing strengths functioned within the larger system of the conversational protocol to promote or inhibit productive discussion amongst mathematics teachers. Additionally, we plan to examine whether the strengths participants name and notice - and thus participants' professional vision - evolved over the two-year program.

## References

Cohen, E. G., & Lotan, R. A. (2014). *Designing groupwork: Strategies for the heterogeneous classroom* (3rd ed.). Teachers College Press.

Coles, A. (2019). Facilitating the use of video with teachers of mathematics: Learning from staying with the detail. *International Journal of STEM Education*, *6*(1), 5. https://doi.org/10.1186/s40594-018-0155-y

Gaudin, C., & Chaliès, S. (2015). Video viewing in teacher education and professional development: A literature review. *Educational Research Review*, *16*, 41–67. https://doi.org/10.1016/j.edurev.2015.06.001

Goodwin, C. (1994). Professional Vision. *American Anthropologist*, *96*(3), 606–633.

Jilk, L. M. (2016). Supporting teacher noticing of students' mathematical strengths. *Mathematics Teacher Educator*, *4*(2), 188–199. https://doi.org/10.5951/mathteaceduc.4.2.0188

Karsenty, R., Peretz, Y., & Heyd-Metzuyanim, E. (2019). *From judgmental evaluations to productive conversations: Mathematics teachers' shifts in communication within a video club*. *TWG18*(34). https://hal.science/hal-02422561

Louie, N. L. (2018). Culture and ideology in mathematics teacher noticing. *Educational Studies in Mathematics*, *97*(1), 55–69. https://doi.org/10.1007/s10649-017-9775-2

McDonald, J. P., Mohr, N., Dichter, A., & McDonald, E. C. (2007). *The power of protocols: An educator's guide to better practice* (2nd ed.). Teachers College Press.

Santagata, R., König, J., Scheiner, T., Nguyen, H., Adleff, A.-K., Yang, X., & Kaiser, G. (2021). Mathematics teacher learning to notice: A systematic review of studies of video-based programs. *ZDM – Mathematics Education*, *53*(1), 119–134. https://doi.org/10.1007/s11858-020-01216-z

Schwarts, G., Karsenty, R., & Arcavi, A. (2022). Ignoring, upholding, redirecting, provoking: Ways of enacting norms in a video-based professional development. *Twelfth Congress of the European Society for Research in Mathematics Education (CERME12)*, *TWG27*(11). https://hal.science/hal-03746283

van Es, E. A. (2011). A framework for learning to notice student thinking. In *Mathematics Teacher Noticing*. Routledge.

van Es, E. A., & Sherin, M. G. (2008). Mathematics teachers' "learning to notice" in the context of a video club. *Teaching and Teacher Education*, *24*(2), 244–276. https://doi.org/10.1016/j.tate.2006.11.005

# THE INFLUENCE OF BILINGUALISM ON CHILDREN'S SELF-EFFICACY BELIEFS IN MATHEMATICS

Yuhwa Hong

University of Massachusetts Amherst

*This study investigated the association between bilingualism and children's self-efficacy beliefs in mathematics using fourth-grade U.S. data from the 2019 TIMSS. Employing the Students Confident in Mathematics (SCM) scale in TIMSS as a dependent variable and including control variables such as gender, academic achievement, engagement, and socioeconomic status, this study showed that bilingual children have significantly higher self-efficacy beliefs in mathematics than non-bilingual children.*

According to Bandura (1986), self-efficacy beliefs determine individuals' thoughts, feelings, motivation, and behaviors. One's beliefs about their capacity to succeed in their endeavors have a strong impact on their subsequent successes or failures (Bandura, 1986). Previous studies in the field of education have convincingly demonstrated the importance of self-efficacy beliefs in students' learning and subsequent academic performance through their influence on the choices they make, effort they invest into their actions, persistence they exercise, and anxiety they experience (Pajares & Schunk, 2001; Mackay & Parkinson, 2010; Schunk, 1982). There is evidence that, along with the children's actual capabilities, their perception of such capabilities plays a significant role in performance and achievement. If students should have "both the 'will' and the 'skill' to be successful in classrooms" (Pintrich & De Groot, 1990, p. 28), the 'will' can be constructed from an individual's beliefs in their capacity to be successful. Accordingly, beliefs are also strongly related to course selection and career choice and are among the factors that explain the gender bias in science, technology, engineering, and mathematics (STEM) education (Webb-Williams, 2018). Therefore, exploring which factors can influence the development of self-efficacy beliefs helps understand one's academic and/or non-academic life trajectories.

This paper explores the effect of bilingualism on children's self-efficacy beliefs in mathematics, considering it as a potential factor with a significant role in shaping self-efficacy beliefs. Utilizing data from the 2019 Trends in International Mathematics and Science Study (TIMSS), the study aims to shed light on how diverse cultural circumstances and experiences, relating to various language usages and belonging to different communities, may influence the way children interpret the sources of self-efficacy beliefs and perceive their academic abilities in mathematics.

## THEORETICAL FRAMEWORK

From a traditional perspective, self-efficacy beliefs are often viewed as internal mental representations (Gee, 2008). However, Bandura's social cognitive theory also



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 57-64). PME.

recognizes the significant impact of external factors on the development of these beliefs. The traditional assumption that individuals internalize information uniformly, regardless of interpersonal differences, contradicts observed disparities in children's learning and thinking (Gee, 2008). This highlights the importance of sociocultural dimensions, emphasizing the role of individual, social, and cultural factors in shaping self-efficacy beliefs.

When examining children's self-efficacy through a sociocultural lens, language should emerge as a crucial factor in understanding how these beliefs take shape within their social and cultural circumstances. Language serves as the cornerstone of an individual's conceptual ecology, facilitating cognitive growth and mediating higher-order thinking (John & Brader-Araje, 2002).

The importance of language makes it meaningful to investigate self-efficacy development in bilingual children. Acquiring more than one language demands increased effort, perseverance, and flexibility. According to Bandura (1997), these qualities play a major role in constructing positive self-efficacy beliefs. Beyond the traditional perspective, adopting a sociocultural lens becomes crucial, viewing bilingual children as participants in multiple communities rather than just individuals proficient in multiple languages (Moschkovich, 2002). Therefore, examining the influence of bilingualism on self-efficacy beliefs is essential for unraveling the intricate relationship among languages, communities, cultures, relationships, and experiences, broadening our perspective on the development of strong self-efficacy in education.

## LITERATURE

Research on students' self-efficacy has predominantly focused on high school and college-aged individuals in predominantly White settings. However, some studies suggest that contextual and demographic factors, such as gender, ethnic background, and learning domain, significantly impact outcomes in this domain (Usher, 2009; Britner & Pajares, 2006; Lent, Lopez, & Bieschke, 1991; Usher & Pajares, 2006). While a limited amount of research has explored the influence of contextual and demographic factors on efficacy beliefs, it is imperative for researchers to place greater emphasis on cultural dimensions that may significantly shape self-efficacy beliefs in diverse settings (Klanssen, 2004).

Examining the notion that efficacy beliefs can vary across cultures, Klanssen's (2004) study on 270 Grade 7 students (Indo Canadian and Anglo Canadian) found that Indo Canadian students displayed a more hierarchical orientation in mathematics efficacy beliefs. Social comparison significantly influenced their motivation and efficacy beliefs. Although the study did not explicitly address the cultural significance of language, participants were categorized into two distinct language groups, potentially indicating language as a factor in understanding cultural influences.

In Clifton-Sprigg's (2015) exploration of the impact of bilingualism at home on early childhood cognitive and non-cognitive performance using Scottish Government data, findings revealed that children displayed comparable skills, regardless of language spoken. Bilingual children performed similarly to monolingual peers, even in the English Vocabulary Naming Exercise. The study underscored heterogeneity within bilingual families, indicating that factors like parental background might contribute to variations in children's cognitive and non-cognitive skills development.

In addition to the previous studies, this research aims to predict the relationship between bilingualism and students' self-efficacy beliefs in mathematics, considering other control factors that might also influence self-efficacy beliefs. The specific research questions are as follows: (1) What is the relationship between being bilingual and children's self-efficacy beliefs in mathematics? (2) What other factors contribute to children's self-efficacy beliefs in mathematics?

## DATA

I used data from the TIMSS, a series of large-scale international assessments of mathematics and science skills. The TIMSS is designed to provide cross-national data on fourth- and eighth-grade students' achievements and various information around students' achievements in mathematics and science.

This study focuses on fourth-grade students in the United States who participated in the 2019 TIMSS. The choice of a single-country sample is attributed to the diverse reasons internationally for children speaking different languages at home and school. The decision to concentrate on the USA ensures a focused examination of specific reasons and circumstances within a single country.

After restricting the sample this way, it transpired that nearly 20% of cases had missing data for the variables used in this study, typically a school-level variable. I concluded that data were missing at random based on missing data analyses and used multiple imputation to address this missingness (Manly & Wells, 2015). All variables were included in the imputation models, as well as the primary sampling unit, strata variables, and the appropriate weights.

## VARIABLES

The study used the Students Confident in Mathematics (SCM) scale in the TIMSS as the dependent variable to measure children's self-efficacy beliefs. Emphasizing a sociocultural perspective, the SCM aligns with the study's purpose, focusing on the experiences of bilingual children and families rather than predicting future outcomes. In contrast, PISA's task-specific instruments in Table 1, emphasizing problem-solving potential, differ from SCM, which gauges students' feelings and judgments about mathematics abilities. PISA's scale may be inappropriate for young learners, as they may not make clear self-judgments at each learning step. Therefore, the SCM scale is more suitable, measuring confidence in subject-specific concepts without specific problem assessments.

Bilingualism, the main dependent variable, was determined by students' use of the test language at home. In America, where English is the test language, responses to the question "How often do you speak English at home?" categorized students as non-bilingual (always/almost always) or bilingual (sometimes/never). Of the sample, 23% were bilingual.

Control variables, such as gender, academic achievements, engagement, and individual and community socioeconomic status, were chosen based on prior research. (Dai & Rinn, 2008; Marsh et al., 2008; Huang, 2013; Schunk, 1984; Sökmen, 2021; Pajares & Schunk, 2001; Pajares, 2005; Tellhed et al., 2017). All these variables were derived from TIMSS students' responses.

| TIMSS | PISA |
|---|---|
| Students Confident in Mathematics<br>How well students think they can do mathematics | Mathematics Self-efficacy<br>How confident students are in their ability to solve mathematics problems |
| 1 Mathematics is more difficult for me than for many of my classmates. | 1 Using a train timetable to work out how long it would take to get from one place to another. |
| 2 I usually do well in mathematics. | 2 Calculating how much cheaper a TV would be after a 30% discount. |
| 3 Mathematics is not one of my strengths. | 3 Calculating how many square meters of tiles would be needed to cover a floor. |
| 4 I learn things quickly in mathematics. | 4 Calculating the petrol-consumption rate of a car. |
| 5 Mathematics makes me nervous. | 5 Understanding graphs presented in newspapers. |
| 6 I am good at working out difficult mathematics problems. | 6 Finding the actual distance between two places on a map with a 1:10 000 scale. |
| 7 My teacher tells me I am good at mathematics. | 7 Solving equations like 3x+5=17 and 2(x+3) =(x+3)(x–3). |
| 8 Mathematics is harder for me than any other subject. | |
| 9 Mathematics makes me confused. | |

Table 1: Instruments in TIMSS and PISA. SOURCE: IEA’s Trends in International Mathematics and Science Study, TIMSS 2019 and OECD, 2015 PISA in Focus.

**METHODS**

The study initially compared means and standard deviations of variables between bilingual and non-bilingual students, revealing potential disparities in mathematics self-efficacy. Subsequently, ordinary least squares (OLS) regression predicted students' math self-efficacy beliefs using the SCM scale as the dependent variable, incorporating bilingualism and five control variables.

## LIMITATIONS

This study acknowledges limitations in data and methods, specifically the unavailability of student-level socioeconomic status data due to the absence of parental questionnaire results in America. To compensate, the number of books at home was used as a proxy for individual socioeconomic status. While unable to confirm the direct association's significance between a student's actual socioeconomic status and the number of books at home, combining this with school SES aimed to collectively control for individual and community factors (Güven, 2019).

## RESULTS

Descriptive analysis results, detailed in Table 2, highlight differences between bilingual and non-bilingual children. A significant gap exists in achievement scores, with bilingual students averaging 24.64 points lower than non-bilingual peers. However, there is no statistically significant difference in self-efficacy beliefs. Bilingual children often come from socioeconomically disadvantaged backgrounds; a trend mirrored in schools' SES. The variable of interactions with mathematics teachers, reflecting children's engagement, shows no statistically significant difference between bilingual and non-bilingual students.

| | All Children | | Bilingual Children | | Non-Bilingual Children | | Difference |
|---|---|---|---|---|---|---|---|
| | Mean | Std. | Mean | Std. | Mean | Std. | |
| Achievement | 534.49 | (1.89) | 515.13 | (2.52) | 539.77 | (2.07) | −24.64*** |
| Self-Efficacy Beliefs | 9.96 | (0.02) | 9.95 | (0.05) | 9.97 | (0.03) | −0.02 |
| Female | 0.49 | (0.01) | 0.53 | (0.01) | 0.48 | (0.01) | 0.052*** |
| Number of Books at Home | | | | | | | |
| 1 (0–10) | 0.17 | (0.01) | 0.21 | (0.01) | 0.15 | (0.01) | 0.06*** |
| 2 (11–25) | 0.24 | (0) | 0.3 | (0.01) | 0.22 | (0.01) | 0.09*** |
| 3 (26–100) | 0.32 | (0) | 0.28 | (0.01) | 0.32 | (0.01) | −0.04** |
| 4 (More than 100) | 0.28 | (0.01) | 0.2 | (0.01) | 0.3 | (0.01) | −0.1*** |
| School SES | | | | | | | |
| Disadvantage | 0.55 | (0.02) | 0.66 | (0.04) | 0.53 | (0.02) | 0.13*** |
| Middle | 0.21 | (0.02) | 0.17 | (0.03) | 0.22 | (0.02) | −0.05*** |
| Affluent | 0.24 | (0.02) | 0.17 | (0.02) | 0.25 | (0.02) | −0.08*** |
| Interactions | 3.99 | (0.02) | 3.97 | (0.03) | 3.99 | (0.02) | −0.03 |
| Observations | 10115 | | 2333 | | 7782 | | |

Table 2. Means and standard errors of the estimates, for all variables

The results in Table 3 confirm a statistically significant positive correlation between bilingualism and children's self-efficacy in mathematics. Model 2 reveals significant associations between gender, achievements, interactions with math teachers, and school SES with children's self-efficacy in mathematics, aligning with existing research.

| | Model 1 | | Model 2 | |
|---|---|---|---|---|
| | Coef. | Std. | Coef. | Std. |
| Bilingual | −0.02 | (0.05) | 0.29*** | (0.04) |
| Female | | | − 0.31*** | (0.05) |
| Achievement | | | 0.01*** | (0.00) |
| Interactions with Math Teachers | | | 0.12*** | (0.02) |
| Number of Books at Home | | | 0.06*** | (0.02) |
| School SES | | | −0.20*** | (0.03) |

Table 3. Results of OLS Regression

This positive relationship is reflected in the predicted self-efficacy values for bilingual and non-bilingual children, which are 10.19 and 9.90, respectively. Figure 1 further illustrates the difference in predictive margins between non-bilingual and bilingual students, with 95% confidence intervals confirming the statistically significant higher belief in self-efficacy among bilingual students.

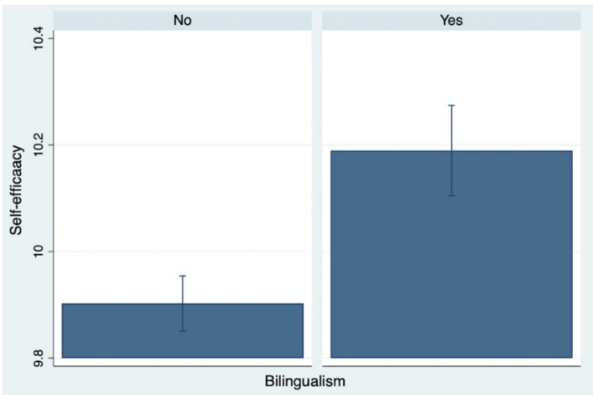



Figure 1. Predictive Margins between Bilingual vs. Non-Bilingual

## BILINGUALISM, ACHIEVEMENT, AND SELF-EFFICACY

Descriptive analyses indicate that bilingual children in America are statistically more likely to come from disadvantaged backgrounds, with significantly lower achievement scores than non-bilingual peers. Given the paramount influence of prior achievement on strong self-efficacy development in mathematics, it is logical to anticipate a negative impact on bilingual children due to their statistically lower achievements compared to non-bilingual children.

However, the results of the OLS analyses suggest an underlying aspect beyond the apparently lower self-efficacy beliefs in bilingual children. The relationship between bilingualism and children's self-efficacy beliefs in mathematics initially appears negative, suggesting that bilingual children might have lower self-efficacy beliefs. Yet, it is not statistically significant, preventing us from drawing conclusions about the low levels of belief in self-efficacy among bilingual children. Subsequently, when the achievement variable was introduced, the direction of the bilingualism variable coefficient shifted from negative to positive, and the estimates became statistically significant. This implies that the positive direct effect of being bilingual on self-efficacy is not evident as an overall effect without controlling for achievement. The indirect effect of being bilingual on self-efficacy through achievement masks the direct effect, given the negative relationship between being bilingual and achievement.

In summary, without controlling for children's achievements, self-efficacy in bilingual and non-bilingual children does not differ. However, isolating bilingualism by controlling for other factors, particularly achievement scores, reveals a significant difference: bilingual children exhibit notably higher self-efficacy beliefs in mathematics compared to their non-bilingual students.

## CONCLUSION AND FURTHER STUDIES

The results of this study are highly intriguing. Yet, what proves more meaningful are the new questions and directions for further research that this study elucidates. Through this study's analyses, I cannot examine the causations and the role of achievements as mediators among the three constructs. Therefore, in future studies, path analysis can be conducted to explore their relationships in-depth and answer whether the positive influence of bilingualism can overcome the negative influence of low achievements.

Additionally, TIMSS includes data for both eighth and fourth grades. This allows me to examine whether there are any movements and changes in the influences of being bilingual and the effect of achievements on students' self-efficacy beliefs in mathematics across different age groups. While the levels of mathematics they learn and their language fluency are expected to change as bilingual children advance to higher grades, the two grade-based populations might exhibit specific differences.

Moreover, as the next step in this quantitative study, a qualitative research phase can be undertaken to explore how and why being bilingual has a positive effect on children's levels of belief in their self-efficacy in mathematics, based on Bandura's social theory and sociocultural theory. This qualitative research can provide insights into the environmental, situational, and contextual influences of bilingualism related to family circumstances, experiences in schools, communities, and societies, shaping diverse interpretations of the sources of self-efficacy beliefs in mathematics.

## References

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, *84*(2), 191.

Bandura, A. (1986). Social foundations of thought and action. *Englewood Cliffs, NJ*, *1986*(23-28).

Britner, S. L., & Pajares, F. (2006). Sources of science self-efficacy beliefs of middle school students. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, *43*(5), 485-499.

Clifton-Sprigg, J. (2015). Bilingual children lag behind in language learning early on, but catch up by age five.

Gee, J. P. (2008). A sociocultural perspective on opportunity to learn. *Assessment, equity, and opportunity to learn*, 76-108.

Güven, U. (2019). The effect of socioeconomic status on student achievement: A TIMSS study. *Contemporary approaches in education and social science*, 89-99.

Jones, M. G., & Brader-Araje, L. (2002). The impact of constructivism on education: Language, discourse, and meaning. *American Communication Journal*, *5*(3), 1-10.

Klassen, R. M. (2004). A cross-cultural investigation of the efficacy beliefs of South Asian immigrant and Anglo Canadian nonimmigrant early adolescents. *Journal of Educational Psychology*, *96*(4), 731.

Lent, R. W., Lopez, F. G., & Bieschke, K. J. (1991). Mathematics self-efficacy: Sources and relation to science-based career choice. *Journal of counseling psychology*, *38*(4), 424.

Mackay, J., & Parkinson, J. (2010). Gender, self-efficacy and achievement among South African Technology teacher trainees. *Gender and Education,* 22(1), 87-103.

Manly, C. A., & Wells, R. S. (2015). Reporting the use of multiple imputation for missing data in higher education research. *Research in Higher Education*, *56*, 397-409.

Pajares, F., & Schunk, D. H. (2001). Self-beliefs and school success: Self-efficacy, self-concept, and school achievement. *Perception*, *11*(2), 239-266.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of educational psychology*, *82*(1), 33.

Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes.* Harvard University Press.

Schunk, D. H. (1982). Effects of effort attributional feedback on children's perceived self-efficacy and achievement. *Journal of educational psychology*, *74*(4), 548.

Usher, E. L., & Pajares, F. (2006). Sources of academic and self-regulatory efficacy beliefs of entering middle school students. *Contemporary educational psychology*, *31*(2), 125-141.

Usher, E. L. (2009). Sources of middle school students' self-efficacy in mathematics: A qualitative investigation. *American Educational Research Journal*, *46*(1), 275-314.

# RELEVANT MEASUREMENT SKILLS TO SOLVE WORD PROBLEMS WITH LENGTHS

Jessica Hoth and Constanze Schadl

Goethe-University of Frankfurt, Friedrich Schiller University Jena (Germany)

*Several instances in daily life require dealing with lengths. These challenges are generally connected to real-life situations and require (among other things) skills to measure, estimate, or convert lengths. In order to analyse the extent of these interrelations, we assessed 277 third and fourth-grade students' skills in solving word problems with lengths. We used a latent multiple regression model to explore the predictive contributions of length measurement, estimation, and conversion skills. Even though all latent variables are significantly correlated, only students' length conversion skills explain relevant variance in the word problem-solving skills, while their length measurement and estimation skills did not.*

## THEORETICAL BACKGROUND

Measurement plays a central role in our everyday life; therefore, it is an international goal of mathematics education that students can reasonably deal with measures. Focusing on lengths as one specific measure that is already introduced in primary school (of course, students bring prior knowledge about lengths from kindergarten and individual experiences), solving word problems with lengths requires some knowledge and skills. For example, if you climb up a huge building, count your steps, and reach the top after 768 steps, you may ask yourself how high you went up. In order to find out, you may either estimate the height of one step or measure it (we assume it to be 15 cm). Thus, you are located at a height of $768 \cdot 15\ cm = 11{,}520\ cm$. Of course, you can convert the given length to meters. Here, you need to know and use the conversion factor correctly (e.g., 11,520 cm = 115.20 m). The example shows that length measurement, estimation, and conversion skills may be relevant to solve real-world problems (or, in an educational context, word problems).

The term word problems is used here to describe specific mathematics tasks from an educational context where background information on the problem is presented as text (Boonen et al., 2013). These word problems are generally very challenging to students of all ages (Daroczy et al., 2015; Pongsakdi et al., 2020). This may be due to the several processes they must successfully manage to reach an adequate solution. This includes initially understanding the general problem, developing a mathematical model by simplifying and restructuring, solving the model with mathematical tools, interpreting the results with respect to the original problem, validating whether the result is appropriate and reasonable and, finally, communicating the solution (Depaepe et al., 2015). These different steps require different knowledge and skills. As Fung & Swanson (2017) propose, students' working memory becomes relevant in word problem-solving processes. Pongsakdi et al. (2020) showed that text comprehension



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 65-73). PME.

and arithmetic skills are related to students' word problem-solving skills. Strohmaier et al. (2022) contributed that verbal skills consistently predicted students' word problem-solving skills, while arithmetic skills only predicted the correct solution if word problems required calculations. In addition, they found that spatial abilities become relevant if word problems contain visualizations.

If word problems contain lengths (as proposed by the introductory example), additional skills may become relevant that are needed to deal with lengths reasonably. In order to make sense of length measures and to operate with them in the multi-step word problem-solving process, students need a concept of length measurement (Tan-Sisman & Aksu, 2011). This includes several skills such as visual estimation, measuring with rulers, and performing conversions (Tan-Sisman & Aksu, 2011). Some indication already exists that there is an interrelation between understanding "measurement concepts, carrying out operations with measurement, and solving word problems involving measurement" (Tan-Sisman & Aksu, 2012, p. 151). In this approach, we aim to analyse the specific meaning of these skills for solving word problems with lengths. Thus, we focus on exploring the following two hypotheses:

H1: Solving word problems with lengths significantly correlates with length measurement, estimation, and conversion skills.

H2: Students' length measurement, estimation, and conversion skills show predictive contributions to students' skills to solve word problems with lengths.

## METHODOLOGICAL APPROACH

In order to evaluate these two hypotheses, we developed paper and pencil tests for each of the proposed parts (word problems, measurement, estimation, and conversion) to assess students' corresponding skills. The test on word problems with lengths held nine short word problems (about 2-4 sentences) and asked students to solve situations with lengths. We systematically varied the operation underlying each task (two tasks required addition, two tasks subtraction, two tasks multiplication, another two tasks division, and, finally, one task focused on proportional relations). Some tasks were given in an open response format, and the other part was given in a single choice format, asking the students to choose the right one of four approaches with the corresponding solution. Half of the sample received tasks 1-5 as open response tasks and tasks 6-9 as single choice, while the other half received tasks 1-5 as single choice and tasks 6-9 as open response. The combination of open response and single choice format was chosen to simplify the students' requirements by presenting them with four possible solutions to choose from but, also, asking them to create a solution themselves—as generally typical for word problems. Figure 1 shows an item example. On the left side of the figure, the item is presented in its open response format; on the right side is the single choice version of the corresponding item.

To assess students' length measurement skills, a total of 17 items were developed in a paper and pencil test. Eight items (M1) required the students to measure the length of a given line with a regular ruler (30 cm) that was either straight (six items) or composed of different lines (three items). Five items (M2) required the students to draw a straight line to a given length with a regular ruler (30 cm). Finally, another four items (M3) asked the students to measure the length of a straight line with a broken ruler (0 and 1 cut off). Figure 2a shows one item from each of the three item blocks about measurement.

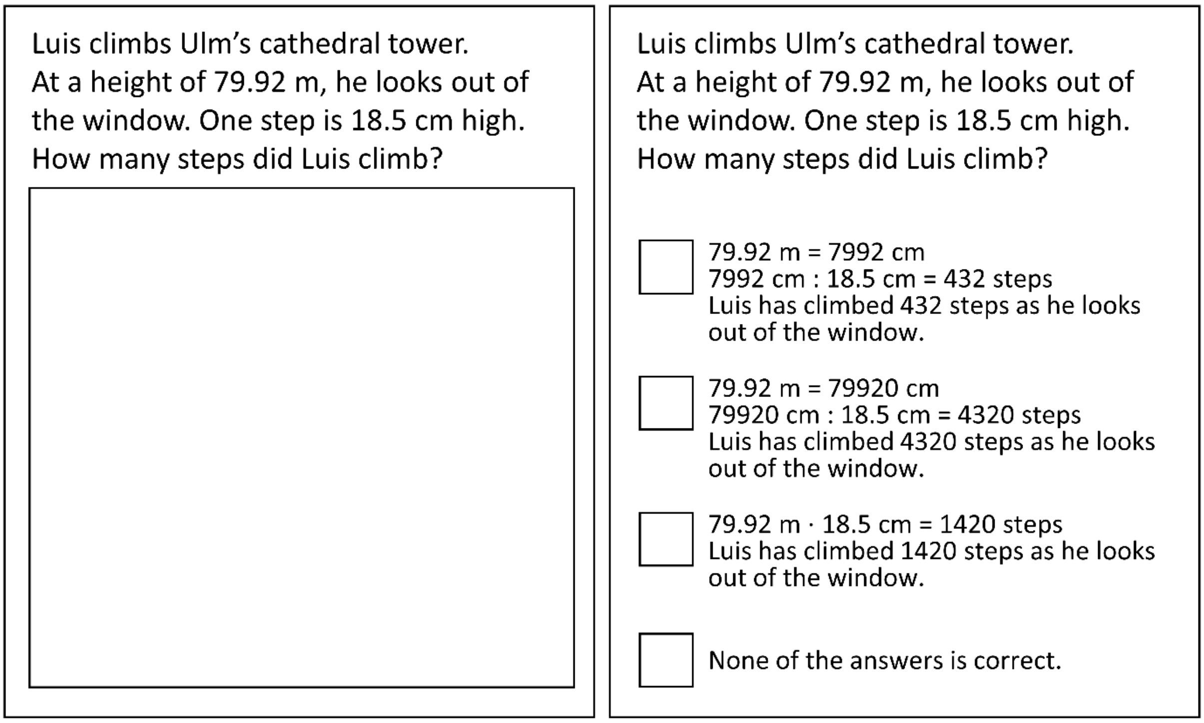



Figure 1: Word problem with lengths in two formats (open response & single choice) as an item example

The test on students' length estimation skills held 12 items. As proposed by previous research (Heinze et al., 2018; Hoth et al., 2022), we varied the objects' size and accessibility to validly assess all facets of students' length estimation skills. In six of these items (E1), students were either asked to estimate the length of small (< 12 cm) and touchable lines or draw lines to a given length < 12 cm (lines are also touchable after their construction). Seven items (E2) held objects that were not small (> 12 cm). Figure 2b shows an example item for each of the two parts. All 13 items asked the children to estimate the length of an object in a standardized measure (or draw a line of the length of a given standard measure), providing quantifying and standardized measures.

| | | |
|---|---|---|
| M1: How long is the line you see here? Measure the lines using the ruler.<br><br>The line is ____ cm long. | M2: Use your ruler and draw a line that is 13 cm long. | M3: How long is the line you see here? Measure the lines using the broken ruler.<br><br>The line is ____ cm long. |

Figure 2a: Example items for the length measurement test

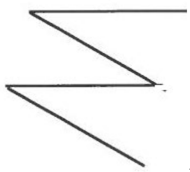
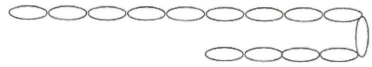
| | |
|---|---|
| E1: How long is line B approximately?<br><br>Line B<br>Line B is about ______ cm long. | E2: How long is this chain approximately?<br><br>The chain is about ____ cm long. |

Figure 2b: Example items for the length estimation test

| | | | |
|---|---|---|---|
| C1: Convert.<br>170 mm = ____ cm | C2: Convert.<br>24 km = ____ m | C3: Convert.<br>807 cm = __m __cm | C4: Larger smaller or equal? Enter the correct sign.<br><br>10 m ☐ 100 cm |

Figure 2c: Example items for the length conversion test

The test on students' conversion skills consisted of 16 items altogether (see Figure 2c for exemplary items). Four items (C1) are required to convert a given length from a smaller unit to the next larger unit. Another four items (C2) required the students to convert a given length from a larger unit to the next smaller one. Again, in another four items (C3), the students were asked to convert a given length in a mixed representation of units. Finally, the last four items (C4) required the students to decide which one of two given lengths is smaller/larger or if both expressions are of equal lengths.

The internal consistency of all independent variables was satisfactory with $.65 < \alpha < .72$. Expert judgements from researchers and primary school mathematics teachers ensured the validity of all test parts.

A total of $N = 277$ students from 15 German third and fourth grade classes and four different schools in and around Frankfurt, Germany, participated in the study (12% third-grade students). 47.3% of this total sample was female, 50.8% was male, and 1.8% was various, and in one case, the gender was not specified. Other demographic variables and personal data were not assessed.

The students solved the 54 items in a paper and pencil format in their classroom. Each of the four test parts (measurement, estimation, conversion, and word problems) was time-limited. The time limit was piloted in advance in order to ensure that students had enough time to solve each of the corresponding tasks. However, to ensure that no test part was omitted due to a lack of time, students were instructed to work on only one test part in a given time. In addition, each test part started with a discussion of one example task. Here, the students could ask questions if they needed help understanding the general task format. After this discussion, the testing time started. In order to ensure equal testing conditions in each class, a trained teacher administered the test using standardized test manuals. The overall testing time was 60 minutes. The students were allowed to take a short break after half of the time.

The data was scored dichotomously (0 = missing or incorrect response, 1 = correct response). Since the time limits ensured that students had enough time to solve each task and each test part, missing responses were interpreted as missing ability. The data for the length estimation test was scored regarding the deviation between the student's estimate and the actual length of the to-be-estimated object. If the student's estimate deviated by not more than 10%, the answers were scored with 3 points, a deviation of $10\% < x \leq 25\%$ was scored as 2 points, estimates deviating between $25\% < x \leq 50\%$ were scored with 1 point and estimates deviating by more than 50%, were scored as incorrect (0).

| | Measurement | Estimation | Conversion | Word problems |
|---|---|---|---|---|
| Measurement | 1 | | | |
| Estimation | .67*** | 1 | | |
| Conversion | .74*** | .65*** | 1 | |
| Word Problems | .36** | .45*** | .61*** | 1 |

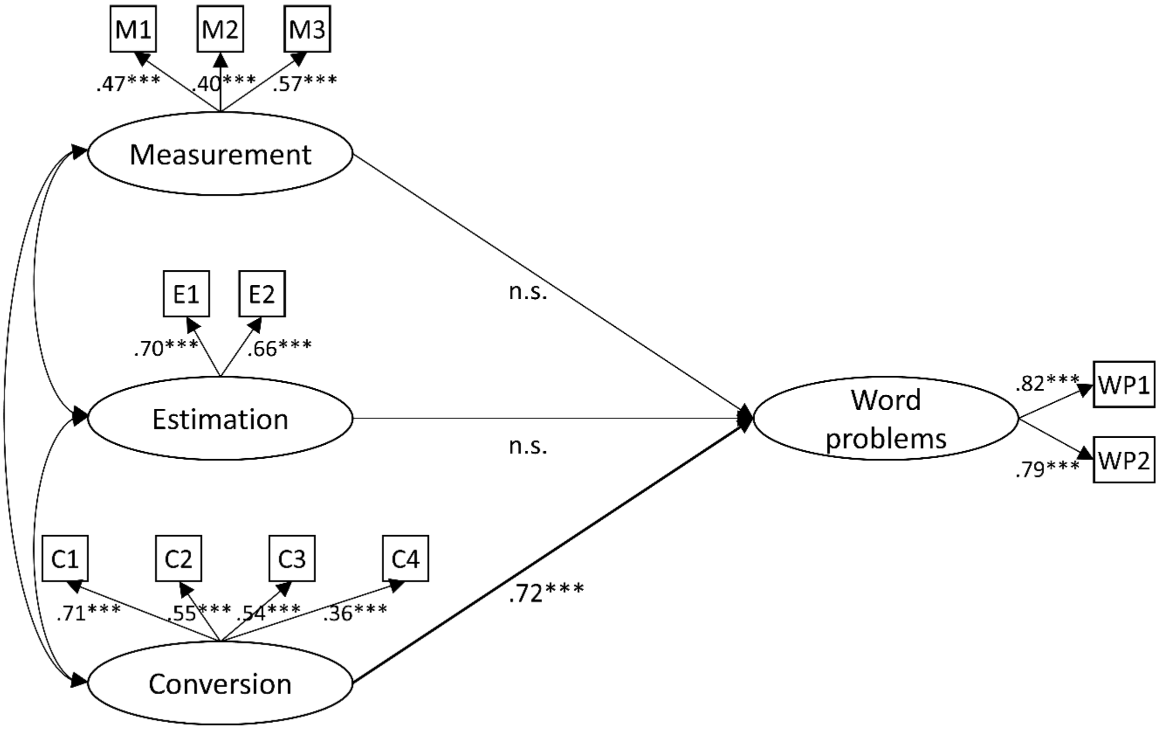
Table 1: Correlation matrix of the latent variables (**$p < .01$. ***$p < .001$)

In order to analyse the interrelations between students' skills to solve word problems with lengths and their skills to measure, estimate and convert lengths, we modelled four latent variables based on the students' scores in each of the four test parts. We analysed their correlation in order to get an indication of their relations (H1). To evaluate the second hypothesis (H2), we specified a latent multiple regression model with the latent variable specifying the skill to solve word problems with lengths as the dependent variable. All analyses were conducted using MPlus (version 5).

## RESULTS

Table 1 shows the correlation matrix for the four specified latent variables. All variables correlate highly significantly with correlation coefficients between .36 and .74, indicating that all skills are interrelated.

Therefore, the first hypothesis (H1) can be verified. Focusing on H2, Figure 3 shows the model and its resulting interrelations.




*Note.* M is the abbreviation for measurement, E for estimation, C for conversion, and WP for word problems.

Figure 3: Latent multiple regression model (H2)

It shows that the variance in students' skills to solve word problems is significantly predicted by their skills to convert lengths, while their skills to measure and estimate lengths cannot explain any additional variance. As proposed in Figure 3, the model explains 41% of the variance in students' skills to solve word problems with lengths. Thus, 59% of the variance in students' word problem-solving skills may be explained by other factors that were not considered in this model such as working memory or arithmetic or language skills as proposed by previous studies (Fung & Swanson, 2017; Strohmaier et al., 2022). The model's fit was acceptable ($\chi^2 = 71.52$, df = 38, $p < 0.01$, CFI = 0.94, TLI = 0.91, RMSEA = 0.056, SRMR = 0.053).

## DISCUSSION

Dealing with lengths is relevant in many situations of daily life. To prepare students for these daily requirements, they are generally presented with word problems in their mathematics education that describe a specific problem involving lengths, such as the example presented before: After climbing up a building and counting the steps, you could determine the height that you are located at if you determine the height of one step. Here, you could measure or estimate, for example. In addition, you may use conversion to resolve the correct size. In this regard, we analysed the predictive contributions of these three proposed skills (measurement, estimation, and conversion) for students' word problem-solving skills. Analysing a latent correlation matrix based on data from 277 third and fourth-grade students indicates that all skills correlate significantly with each other, suggesting that students' skills to solve real-world problems with lengths are connected with their skills to measure, estimate, and convert. However, regarding multivariate relations, a latent multiple regression model shows that only students' length conversion skills predict their word problem-solving skills, while measurement and estimation skills cannot explain any additional variance. This suggests that in many situations with lengths, converting a given length into another unit may be more relevant for a sufficient solution than length measurement or estimation skills.

Of course, this may be due to the word problems that were part of the test instrument. One limitation of this study is that we could only present the students with a limited amount of word problems because these kinds of tasks require much time and are a major cognitive challenge for many children in primary school. We focused on varying the operations needed in each word problem but did not explicitly focus on varying the estimation or measurement needed. Of course, there may be word problems that require length estimation and/or measurement skills more specifically than the tasks in our test. Another limitation is that in the current state of data assessment, the sample holds only a few third-grade students (12% of the overall sample). This group of students is—on the one hand—underrepresented in the sample but—on the other hand——explicitly challenged by many of the given requirements because length conversion is just introduced in third grade as well as some of the standard units such as mm and km. However, the data was assessed at the end of the school year—the third graders in the sample should have been introduced to the relevant content. Finally, there is still 59% of variance that the three independent variables cannot explain. As proposed by other research, word problems (naturally also those dealing with lengths) require several other skills, such as language or arithmetic skills (e.g., Strohmaier et al., 2022). There may be some variance in students' word problem-solving skills that these additional factors may explain.

Therefore, further research is needed to clarify which predictive contributions fall on specific skills to deal with lengths when solving word problems (with lengths) compared to other mathematics-specific skills such as arithmetic skills and language skills or general cognitive abilities. Regarding mathematics education, the results indicate that unit conversion skills are especially needed for this kind of word problem that was focused on here. Therefore, teachers must be aware of the requirements that specific problems set for students. As length conversion is generally discussed in 3$^{rd}$ and 4$^{th}$-grade primary school classes, word problems with lengths should be introduced in this awareness.

## References

Boonen, A., Schoot, M., Van Wesel, F., Vries, M. & Jolles, J. (2013). What underlies successful word problem solving? A path analysis in sixth grade students. *Contemporary Educational Psychology, 38,* 271-279. https://doi.org/10.1016/j.cedpsych.2013.05.001

Daroczy, G., Wolska, M., Meurers, W.D. & Nuerk, H.-C. (2015). Word problems: a review of linguistic and numerical factors contributing to their difficulty. *Frontiers in Psychology, 6*, 348. https://doi.org/10.3389/fpsyg.2015.00348

Depaepe, F., De Corte, E., & Verschaffel, L. (2015). Students' non-realistic mathematical modeling as a drawback of teachers' beliefs about and approaches to word problem solving. In B. Pepin & B. Roesken-Winter (Eds.), *From beliefs to dynamic affect systems in mathematics education: Exploring a mosaic of relationships and interactions* (pp. 137-156). Springer International Publishing/Springer Nature. https://doi.org/10.1007/978-3-319-06808-4_7

Fung, W., & Swanson, H.L. (2017). Working memory components that predict word problem solving: Is it merely a function of reading, calculation, and fluid intelligence? Memory & Cognition, 45, 804-823. https://doi.org/10.3758/s13421-017-0697-0

Heinze, A., Weiher, D. F., Huang, H-M., & Ruwisch, S. (2018). Which estimation situations are relevant for a valid assessment of measurement estimation skills. In E. Bergqvist, M. Österholm, C. Granberg, & L. Sumpter (Eds.), *Proceedings of the 42nd Conference of the International Group for the Psychology of Mathematics Education*. (Vol. 3, pp. 67-74). PME.

Hoth, J., Heinze, A., Huang, H.-M., Niedermeyer, I., Weiher, D., & Ruwisch, S. (2022). Elementary school students' length estimation skills - analyzing a multidimensional construct in a cross-country study. *International Journal of Science and Mathematics Education, 21,* 1841-1864. https://doi.org/10.1007/s10763-022-10323-0

Pongsakdi, N., Kajamies, A., Veermans, K. *et al.* (2020). What makes mathematical word problem solving challenging? Exploring the roles of word problem characteristics, text comprehension, and arithmetic skills. *ZDM Mathematics Education, 52*, 33-44. https://doi.org/10.1007/s11858-019-01118-9

Strohmaier, A.R., Reinhold, F., Hofer, S. *et al.* (2022). Different complex word problems require different combinations of cognitive skills. *Educational Studies in Mathematics, 109*, 89-114. https://doi.org/10.1007/s10649-021-10079-4

Tan-Sisman, G., & Aksu, M. (2011). The length measurement in the Turkish mathematics curriculum: Its potential to contribute to students' learning. *International Journal of Science and Mathematics Education*, *10*, 363-385. https://doi.org/10.1007/s10763-011-9304-1

Tan-Sisman, G. & Aksu, M. (2012). Sixth Grade Students' Performance on Length, Area, and Volume Measurement. *Education and Science, 37*, 141-154.

# AN INVESTIGATION OF LENGTH ESTIMATION SKILLS OF HIGH SCHOOL STUDENTS WITH MILD INTELLECTUAL DISABILITY

Hsin-Mei E. Huang[1], Xuan Liao[1], Jessica Hoth[2], Silke Ruwisch[3], Aiso Heinze[4]

[1]University of Taipei, Taiwan, R.O.C., [2]Goethe University of Frankfurt, Germany, [3]Leuphana University of Lueneburg, Germany; [4]IPN-Leibniz Institute of Science and Mathematics Education Keil, Germany

*This study investigated the length estimation skills of high school students with mild intellectual disabilities (N = 39) by means of a paper-and-pencil assessment and interviews. The results showed that the students performed differently in different estimation situations involving size discrepancy and accessibility of to-be-estimated objects. The students tended to underestimate the lengths of daily objects. The uses of body parts, objects in convenience, mental rulers such as 1, 10 and 15 cm as reference points through unit iteration were the strategies reported by the interviewees.*

## INTRODUCTION

Length estimation is one of the practical skills that represent real-life mathematics applications in everyday activities. Furthermore, length estimation competence, which serves as the fundamental basis of spatial measurement, is also a core skill that supports the developments of mathematics and science capabilities and employability (Jones & Taylor, 2009; Tretter & Jones, 2006a, 2006b).

An increasing number of studies have provided information about the length estimation skills of students in general education (normally developing students) across various educational stages (Huang, 2020; Hoth et al., 2023a; Tretter & Jones, 2006a, 2006b). Nevertheless, research concerning the issue among students with mild intellectual disabilities (MID) is relatively scare, particularly, among students with MID above middle school levels. On the one hand, length estimation has been included as crucial content of the mathematics curriculum for basic education across countries (e.g., Andrews et al., 2022; Hoth et al., 2023a), on the other hand, mathematics curricula and instruction provided for high school students in self-contained classes are oriented toward real-life mathematics application and vocational development (Ministry of Education, 2022). Seeing that students at high school level have a lot of experiences in estimating lengths both in-and-out of school activities (Tretter & Jones, 2006a), it is worth exploring the length estimation skills of high school students with MID. More importantly, if teachers have deep understandings of students in such special education apply mathematics knowledge to solve length estimation problems, it may help them make better designs and decisions about the mathematics curriculum and instruction.

Previous studies have suggested that characteristics of estimation situation may influence estimators' ability to estimate lengths such as the size and accessibility of to-be-estimated objects (TBEOs) (Hoth et al., 2023a). Moreover, patterns of estimation



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 73-80). PME.

error frequently occur in students' estimated measurements (Jones et al., 2012). For example, underestimating length measurements is found among students at elementary school (Huang, 2020) and middle school levels (Jones et al., 2012). Nevertheless, the role of the estimation situation in length estimations performed by high school students with MID and what error patterns occur in this group of students remain unclear.

Using effective strategies is a fundamental basis for obtaining a close estimate. The uses of inappropriate strategies, such as guessing without thinking about the reasonableness of the estimate, may lead to over- or under-estimation. Employing body parts or convenient objects as measurement units through unit iteration is frequently found among young students (Huang, 2020). What strategies are used by high school students with MID for estimating the lengths of objects remains unknown.

This study aimed to explore the length estimation skills of high school students with MID (hereafter "students with MID") enrolled in special education self-contained classes. This study included three research questions. 1. To what extent do students with MID perform differently in estimating lengths in different estimation situations? 2. What differences in the frequency of overestimations and underestimations of the estimated lengths can be observed in different estimation situations? 3. What strategies do students with MID use for estimating the lengths of objects?

## THEORETICAL FRAMEWORK

### Mathematical Thinking Involved in Length Estimations

To make a measurement estimate means to determine a quantitative value of an object without the aid of tools. Mathematical thinking involved in measurement estimation includes visual-spatial thinking, concepts of size and scale, and various forms of mental referents constructed from previous experiences of measurement (Joram et al., 1998; Tretter & Jones, 2006b). Accordingly, making a measurement estimation requires knowledge of measurement (e.g., measurement attributes, standard units) and reasoning skill (Jones & Taylor, 2009), rather than merely guessing.

Furthermore, making a judgement on spatial size strongly depends on what an individual has previously perceived and experienced (Joram et al., 1998). Extensive physical measurement experiences of using units for measuring lengths benefits students' understanding of concepts of size and scale. This in turn helps students develop a set of mental reference units (Jones & Taylor, 2009). For developing students' estimation skills, the use of a set of mental reference units for making length estimates has been highlighted for school mathematics (Andrews et al., 2022).

### Estimation Situation and Patterns of Estimation Error in Length Estimation

Previous research has suggested that the characteristics of estimation situations, which involve characteristics of TBEOs and measure units (benchmarks), play an important role in length estimation (Jones et al., 2012). Hoth et al. (2023a) developed and validated a paper-and-pencil assessment using a 3-dimensional model for examining elementary school students' length estimation skills. It is suggested that size and accessibility of TBEOs were the two crucial aspects that may affect students' estimation performance. Moreover, the roles of size and accessibility of TBEOs in length estimation performance revealed in their study were confirmed by another study on the length estimation skills of junior high school students using Hoth et al.'s assessment (paper in preparation).

Although Tretter and Jones (2006a, 2006b) suggested that students across various grade levels were better able to estimate the size of objects within the range of human size, still elementary school students were found to perform differently when estimating lengths of daily objects within 1 meter according to the estimation situations (Hoth et al., 2023a, 2023b; Huang, 2020). Generally, estimating length measurements in a situation in which TBEOs are not accessible seems more challenging for students than one in which the TBEOs are accessible, given the TBEOs are not small (e.g., ≥ 15 centimeters). In contrast, for estimating length measurements in a situation in which TBEOs were small (e.g., ≤ 12 centimeters), regardless whether the TBEOs were accessible or not, elementary school students' performance was superior to the situation involving not small and not accessible TBEOs (Hoth et al., 2023b).

To explore students' estimation errors produced when estimating the lengths of daily-use objects, Jones et al. (2012) found that underestimations seemed to be made frequently in middle school students' estimated answers. Huang (2020) found that a tendency of underestimations occurred in estimating large-sized TBEOs (51-100 centimeters) by Grade 5-6 students. However, a higher frequency of overestimations than underestimations was found for Grade 6 students for the small-sized TBEOs (1-10 centimeters). It is argued that patterns of estimation errors may have resulted from students' preference for using small measure units such as 1 centimeter for processing unit iterations. This in turn easily leads to estimation errors because a complex processing of unit iterations.

**Strategy Used for Estimating Length Measurements**

As noted previously, selecting appropriate measurement units and comparing with the TBEOs based on knowledge of the size of real-life objects is a fundamental approach (Joram et al., 1998). Using measurement units that are familiar to individuals, which may be used as reference points or benchmarks (e.g., body parts, objects, and 1 or 10 centimeters as a mental ruler), through iterating to achieve the estimated answer occurs frequently among elementary school students (Huang, 2020).

Furthermore, estimators tended to select different measure units for estimating measurements depending on the size of TBEOs (Huang, 2020; Joram et al., 1998). For reducing the number of unit iteration that would be needed to be performed, another approach for processing estimation occurs when individuals split the TBEO into small parts before using unit iteration or reference points and then re-compose the parts by performing calculations (Joram et al., 1998).

## Mathematics Achievement of Students with MID

In general, mathematics instruction for students with MID highlights procedural knowledge (e.g., basic arithmetic, number line, and performing mathematical processes involving measurements, Gersib et al., 2024), conceptual knowledge and problem-solving skills (Ministry of Education, 2022) for supporting self-sustainment in daily life and vocational development. Due to moderate limitations in intellectual functioning, communication skills or working memory, students with MID tend to make slow mathematical progress in their understanding of basic numerical skills and quantity-number concepts, compared with normally developing students of similar age (Numminen et al., 2002).

Given the limitations mentioned above, Schnepel et al. (2020) suggested that "prior knowledge seems to be the most important predictor and [is] more important than IQ" (p. 115) for mathematical achievement of elementary school students with MID in basic number skills. Thus, students with MID are capable of performing length estimations with an increase of their experiences of measurement in daily activities.

Considering the above research findings, this study proposed the following hypotheses: Hypothesis 1 is that students with MID perform differently in length estimations depending on the estimation situation. Hypothesis 2 is that differences exist in the frequency of overestimations and underestimations in each estimation situation.

# METHODOLOGY

## Participants, Instruments, and Scoring

The sample consisted of 39 students with MID enrolled in self-contained classes of a special education program (24 males and 15 females) from a high school in a northern city in Taiwan. The participants with MID were from 10$^{th}$ to 12$^{th}$ grade with ages ranging from 15.75 to 18.50 years. All the participants were identified as disabled students by the Ministry of Education in Taiwan.

In the current study, most participants during their elementary to junior high schooling were included in general education settings and learned with adapted curricula oriented toward the mathematics curriculum for general education. Thus, the participants had received instruction on length measurement and estimations in elementary school based on the mathematics curriculum guidelines.

The participants' length estimation competence was measured by the length estimation assessment, which was revised for examining junior high school students' estimation skills in another study (paper in preparation) based on the original instrument used in Hoth et al. (2023a). In the current study, the assessment, which consisted of 25 items, comprised three sections to examine school-aged students' skills in estimating the lengths of daily items with length between 1 millimeter and 1 meter. The three sections, which represented different estimation situations, were categorized by the size of the TBEOs and their accessibility.

The characteristics of the three sections are described briefly as follows. a. Small (S). The S section included nine questions in which the given TBEOs were ≤ 12 centimeters, including seven TBEOs given were touchable and the other two TBEOs were not touchable. For example, a full-size sharpener picture and a 3-mm strip (benchmark) were given in the test booklet for the question “This stripe is 3 mm long. How long is the sharpener on the right side?” b. Not-small and touchable (NST). The NST section included nine questions in which the given TBEOs were physically present and could be touched (e.g., a train picture was given in the test booklet for the question “How long is the train all together?”). c. Not-small and not touchable (NSNT). This section included seven questions involving the TBEOs which were physically present but were not allowed to be touched (e.g., a wastepaper basket was physically presented in front of the classroom without presenting any packages of copy paper for the question “How many packages of copy paper are about as high as the wastepaper basket on the table?”). All the TBEOs given for the NST and NSNT sections were ≥ 15 centimeters.

To collect the data of the participants’ estimation strategies, one-to-one interviews were conducted by asking “What was the strategy that you used for obtaining the estimated answer?” for three questions pertaining to the S (i.e., the sharpener), NST (i.e., the train), and NSNT (i.e., the wastepaper basket), respectively. The interviews were audio taped and transcribed for analysis. The data reported in the study included the responses of six interviewees who were better able to express thinking verbally.

The estimated answers were scored based on their deviation from the actual length of the TBEOs. For answers that did not deviate more than 10% from the actual length, 3 points were given. If the deviation was greater than 10% but less than 25%, 2 points were given. If answers deviated more than 25% but less than 50%, 1 point was given. For answers that deviated greater than 50% and for missing responses, 0 points were allocated. The total possible scores of the S, NST, and NSNT sections were 27, 27, and 21. The total possible score for the assessment as a whole was 75 points.

### Classification of Patterns of Estimation Error and Strategy

For identifying patterns of estimation error, an estimate < -50% of the actual lengths of the TBEOs was defined as an underestimate, whereas an estimate > 50% of the actual lengths of the TBEOs was defined as an overestimate. For each section, the total frequencies and percentage of over- and under-estimations were calculated.

To analyze the strategies used by the interviewees, the uses of reference points as measurement units (e.g., body parts, objects), previous experiences, mental rulers, and guessing were classified based on the data of interviews.

## RESULTS AND DISCUSSION

Considering that the three sections contained different numbers of questions and to account for any missing responses, statistical analyses were executed on the average-per-question score. Table 1 shows the means and standard deviations of the participants’ average-per-question score by section. The score obtained from the NST situation was the highest, followed by the NSNT and then the S situations.

| Section | *M* | *SD* | Over-estimation | | Under-estimation | |
|---|---|---|---|---|---|---|
| | | | *f* | % | *f* | % |
| Small | 0.56 | 0.31 | 50 | 14% (50/351) | 222 | 63% (222/351) |
| Not-small and touchable | 0.89 | 0.54 | 95 | 27% (95/351) | 84 | 23% ( 84/351) |
| Not-small and not touchable | 0.67 | 0.46 | 25 | 9% (25/273) | 135 | 49% (135/273) |

Table 1: The Means and Standard Deviation of the Students' Average-Per-Question Score by Section ($N = 39$)

To examine the effects of estimation situation on the estimation performance of the students, a repeated measures ANOVA was implemented. A Mauchly's test indicated that the assumption of sphericity had been violated, $\chi^2$ (2) = 11.32, $p < .05$. Thus, the degrees of freedom were corrected using Huynh-Feldt estimates of sphericity ($\varepsilon$ = 0.82). The results showed a significant difference between the three sections ($F$ [1.64, 62.33] = 6.34, $p < .01$, $\eta^2 = .14$). Bonferroni corrected post hoc tests revealed that the scores of the NST section were better than those of the S section (Bonferroni test, $p < .001$). No differences in the students' performance were found between the scores obtained in the NSNT and S sections nor between those obtained in the NST and NSNT sections.

With respect to the comparisons of error patterns, Table 1 shows the frequencies and percentage of the error patters in each section. The total frequency of the students' inaccurate answers in the S and NST and NSNT sections were 351, 351, and 273, respectively. For the S and NSNT sections, the results of $\chi^2$ nonparametric tests revealed that the frequencies of underestimation were higher than those of overestimation, $\chi^2$ (1) = 108.77, $p < .001$, and $\chi^2$ (1) = 75.63, $p < .001$, respectively. For the NST section, the result of $\chi^2$ nonparametric test displayed no differences in the frequencies of underestimation and overestimation, $\chi^2$ (1) = 0.68, $p = .41$.

Taking together the findings of the study, the result that the students with MID performed better in the NST than in the S section was in line with Hoth et al.'s (2023b) study. The result that the students obtained similar scores in the NSNT and S sections probably resulted from the characteristics of the two estimation situations, respectively. Concepts of spatial size and reasoning skills are demanded for solving the problems in the NSNT section which highly required the mental use of reference points. If the participants did not have a comprehensive understanding of spatial size and a good construction of mental images of measurement units, such weaknesses may have impaired their success (Tretter & Jones, 2006a, 2006b). This in turn contributed to overestimates and underestimates. Furthermore, knowledge of metric unit conversions between centimeter and millimeter were needed for solving the problems requesting the use of millimeter as unit in the S section. A lack of such knowledge may have led to inaccurate estimated answers and patterns of errors (Tretter & Jones, 2006b). The results showed that no consistent error pattern was found among the three situations. In the study, the participants had the tendency to underestimate the measurements of the TBEOs given in both the S and NSNT sections. The reason for these differences is possibly resulted from poor estimation accuracy on the S and NSNT sections (Jones et al., 2012). However, a further probe on how estimation situation related to error patterns occurs among students with MID is needed. Finally, the results of the study partially supported Hypothesis 1 and Hypothesis 2, respectively.

As to the strategies used by the interviewees, body parts (e.g., the width of a finger), objects (e.g., a grain of rice, the 1-meter ruler shown on the blackboard, the full-sized pictures of the TBEOs or benchmarks), mental rulers (e.g., personal measure units such as 1, 10, and 15 cm) were used as reference points. Most of the interviewees took advantages of the objects presented in the questions and operated unit iteration. Only one interviewee decomposed the train into parts and then recomposed the parts through multiplication. For example, for estimating the length of the train, one interviewee expressed: “I looked at the ruler first, and memorized it, and then went back to the question, and compared the train with the ruler and thought about the approximate length.”

It is noteworthy that some interviewees imagined the length of the 1-meter ruler on the blackboard and guessed that it was 15 centimeter after comparing it with their hand spans. Such inaccurate guesses led to erroneous estimates. Furthermore, the use of previous experience of measurement for making a judgement was also found. For example, for estimating the length of the sharpener, another interviewee indicated: “I measured a pencil sharper in my childhood…um use experiences to make a judgement.” Finally, the metric unit, millimeter, seemed unfamiliar to the interviewees due to its rare use in daily activities. Most of the interviewees had misconceptions about metric unit conversion. They tended to hold an inaccurate conception that “1 cm = 1000 mm.”

## IMPLICATIONS FOR MATHEMATICS EDUCATION

As Gersib et al. (2024) suggested particularly for students with learning disabilities in mathematics: “A robust knowledge of measurement hold significant value in students’ development of mathematical proficiency” (p. 173). To enhancing length estimation skills of high school students with MID, providing sufficient experiences in modelling appropriate language and measurement processes such as object comparisons, and encourage them to construct correct knowledge of metric units and a repertoire of familiar reference points through practice should be taken into account seriously.

### References

Andrews, P., Xenofontos, C., & Sayers, J. (2022). Estimation in the primary mathematics curricula of the United Kingdom: Ambivalent expectations of an essential competence. *International Journal of Mathematical Education in Science and Technology 53*(8), 2199-2225.

Gersib, J. A., Rojo, M., Ghafaghazi, S., Uy, J., & Doabler, C. T. (2024). Measuring success: Integrating number lines into measurement instruction for students with learning disabilities. *Intervention in Social and Clinic, 59*(3), 173-182.

Hoth, J., Heinze, A., Huang, H.-M. E., Weiher, D. F., Niedermeyer, I., & Ruwisch, S. (2023a). Elementary school students' length estimation skills- Analyzing a multidimensional construct in a cross-country study. *International Journal of Science and Mathematics Education, 21*(6), 1841-1864.

Hoth, J., Heinze, A., Huang, H.-M. E., Weiher, D. F. & Ruwisch, S. (2023b). Elementary school students' length estimation Skills: A cross-country study. Paper presented at 2023 Symposium on Learning Materials and Instructional Innovation: Learning Materials in Interdisciplinary Development and STEAM/STEM Education. Taiwan.

Huang, H.-M. E. (2020). Effects of grade-level and object size on students' measurement estimation performance. *EURASIA Journal of Mathematics, Science and Technology Education, 16* (12), 1923-1938.

Jones, M. G., & Taylor, A. R. (2009). Developing a sense of scale: Looking backward. *Journal of Research in Science Teaching, 46*(4), 460-475.

Jones, M. G., Forrester, J. H., Gardner, G. E., Andre, T., & Taylor, A. R. (2012). Students' accuracy of measurement estimation: Context, units, and logical thinking. *School Science and Mathematics, 112*(3), 171-178.

Joram, E., Subrahmanyam, K., & Gelman, R. (1998). Measurement estimation: Learning to map the route from number to quantity and back. *Review of Educational Research, 68*, 413-419.

Ministry of Education (2022). Curriculum guidelines of 12-year basic education for vocational senior high school: Mathematics field in the domain of general subjects. Taiwan. (In Chinese)

Numminen, H., Service, E., & Ruopila, I. (2002). Working memory, intelligence and knowledge based in adult persons with intellectual disability. *Research in Development Disabilities, 23*, 105-118.

Schnepel, S., Krahenmann, H., Dessemontet, R. S., & Opitz, E. M. (2020). The mathematical progress of students with an intellectual disability in inclusive classrooms: results of a longitudinal study. *Mathematics Education Research Journal, 32*, 103-119.

Tretter, T. R., Jones, M. G., Andre, T., Negishi, A., & Minogue, J. (2006a). Conceptual boundaries and distances: Students' and experts' concepts of the scale of scientific phenomena. *Journal of Research in Science Teaching, 43*(3), 282-319.

Tretter, T. R., Jones, M. G., & Minogue, J. (2006b). Accuracy of scale conceptions in science: Mental maneuverings across many orders of spatial magnitude. *Journal of Research in Science Teaching, 43*(10), 1061-1085.

# DIFFERENCES IN MATHEMATICS LEARNERS ACCORDING TO IN-SERVICE AND PRE-SERVICE TEACHERS

Anna Hummel

Leipzig University, Germany

*Discussing which differences in learners are relevant for or in mathematics learning differs greatly depending on who is asked and their stance of observation. This paper provides empirical insight into group discussions among pre-service and in-service elementary school mathematics teachers, discussing differences in mathematics learners based on their experiences in practice. Comparing categories of differences that were made explicit, reveals similarities and divergences between participating groups. Beyond presenting first results on teachers' innate attributions of difference to learners without theoretically operationalizing lines of differences beforehand, a possible link between the meaning of relevance and the shared context where these differences emerge is discussed.*

## INTRODUCTION

Addressing learners' diversity can be perceived as an integral aspect of attaining equity in mathematics education (Rohn, 2013). Discussing which differences in learners are impactful for mathematics learning and what this entails for mathematics teaching, leads to various perspectives in mathematics education research. Addressing students' diversity is not only important but also demanding, raising questions on how teachers can be prepared to meet this challenge. Therefore, comparing dispositions of pre-service teachers (PSTs) and in-service teachers (ISTs) regarding differences in learners mainly focusses on their trajectory in professional development. For example, professional noticing of children's mathematical thinking differs significantly depending on teaching experience, highlighting a growing expertise in attending to children's strategies and interpreting their understandings based on teaching experience (Jacobs et al., 2010). This paper alters the perspective from *how* to deal with learners' differences towards the question, *which* differences are initially framed as relevant when talking about mathematics learners.

## THEORETICAL AND METHODOLOGICAL FRAMEWORK

Many descriptive approaches link attributed differences to varying learning achievement or opportunities, tracing connections in large-scale achievement studies highlighting disparities along social lines of differences such as 'gender' or 'migration background' (Mullis et al., 2020). Prescriptive approaches on the other hand strive to provide conceptual perspectives on accessible mathematics education for all learners despite or on account of their differences such as inclusive mathematics education (Kollosche et al., 2019) or mathematics for socio-political justice (Gutstein, 2006). Aiming to understand or explain teachers' dispositions towards learners' differences and their subsequent acting brings forth research addressing implicit attitudes and



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 81-88). PME.

stereotypes (Denessen et al., 2022) or aspects within teachers structuring their actions such as beliefs (Voss et al., 2013) or noticing (Louie, 2021). The underlying theoretical understandings of ‘difference’ in this research lead to a wide range of possible reference points for mathematics teaching. In this paper, the perspective on differences is altered by making the *experiences of teaching mathematics* itself the reference point for emerging differences without operationalizing categories of difference beforehand as far as possible. Accordingly, differences are framed in a cultural-sociological perspective, seeing them not as inscribed features of learners, but rather as ascriptive person or group-related attributions (Bräu & Schlickum, 2015). Since mathematics teaching is situated in the organizational logic of school systems including constant processes of assessment and comparison of individuals, the term ‘difference’ entails a classificatory character. ‘Difference’ is seen as relational in regard of a tertium comperationis, following a logic of common and special (Walgenbach, 2014). Whatever is seen as common and special is defined by normative (implicit) orders which constitute the field in which differences emerge. The concept of ‘doing difference’ provides the analytical frame to address these (implicit) norms within the process of constructing differences and to conceptualize differences themselves as products of the stated (implicit) norms by framing them as a “meaningful selection of competing categorizations” (Hirschauer, 2014, p. 183). Focusing on the emergence of differences in discussions among PSTs and ISTs shifts the epistemological interest to which differences are explicated when talking about mathematics learning and to the system of relevance incorporated by teachers. As a methodological approach to reconstruct this system of relevance constituting the process of ‘doing difference’, the research project as a whole draws on the ‘Praxeological Sociology of Knowledge’ (Hummel & Reinhold, 2023). However, the presented insights into data from group discussions only focus on the research question, which differences are explicated. Comparing the emergence of addressed aspects among PSTs and ISTs strives to reveal similarities and divergences in participants references to ‘differences’ based on their experience of teaching mathematics.

## DATA COLLECTION AND ANALYSIS

Due to the research interest in explicated differences in mathematics learners according to mathematics teachers, data collection was based on six group discussions, three groups consisting of PSTs (group sizes: 6, 4 and 3 persons) and three groups consisting of ISTs (group sizes: 4, 3 and 2 persons). The group discussions among PSTs (undergraduate students at a German university) were conducted in July 2022 as part of a pilot study for the main study conducted among German ISTs in May 2023. The participants of five group discussions are from Saxony, one group of ISTs is from Baden-Wurttemberg - all referenced schools are located within urban spaces. At the point of data collection, the PSTs had just finished their four-month internship situated at the end of their third year of academic teacher training during which they taught mathematics once weekly in their assigned school. In contrast, work experience among ISTs varied in between 1,5 and 35 years. The aim of the presented qualitative study was to provoke discussions about differences in mathematics learners without specifying what “differences” might entail. Therefore, data analysis aims to identify differences as discussed by participants on an explicit, communicative level, putting data itself and participants subjective meaning as initial points for analysis. In order to initiate a self-dynamic discourse among participants, which allowed their own and collectively shared meaning of differences in relation to mathematics teaching to emerge, very open discussion impulses were provided. Both impulses asked PSTs and ISTs to talk about their everyday lessons in math and about which differences among learners they experienced.

For the presented first step in data analysis, the tool of qualitative content analysis according to Mayring (2014) was chosen. The aim of using qualitative content analysis was to inductively identify and summarize categories of explicated differences during text analysis. Instead of a hypothesis testing approach, the inductive category formation was led by emerging depictions of difference, leading to categories “which are coming from the material itself, not from theoretical considerations” (Mayring, 2014, p. 80). The focus in this step of analysis was to identify which categories of difference emerged without analyzing their attributed implicit meaning or the accompanying negotiation process. Therefore, results will discuss frequency of explicated categories of differences, treating prevalently stated differences as more relevant within the discussions about mathematics learners – further research steps to substantiate or potentially invalidate this perspective will be provided later in the paper.

## EMPIRICAL INSIGHTS

Data analysis of explicated differences among learners resulted in 262 identifiable segments. The inductive code formation resulted in depictions of difference that could be summarized in four categories: *‘performance related differences*’ (approx. 70%), referring to students’ performances while engaging with mathematics learning or tasks consisting of observable behavior and *‘social differences’* (approx. 18%), addressing students’ social identity attributions. If an argumentation was provided, that consisted of assumptions rather than observations, segments were coded in the category *‘affect or cognition related differences’* (approx. 8%). Within the category *‘attributing labels’* (approx. 5 %) attributions to learner’s identities such as “slow mind”, “weak learner” or “front runner” were coded separately in *‘total attributions’* when referring to individual children, or as *‘general attributions’* when concerning the whole class, e.g. describing class composition as “many weak learners, almost no middle field, many/ few strong learners”. The surplus in total percentage is due to the same segments being coded in different categories due to their ambiguity in possible meaning, as addressed. The results provide the depicted distribution and prevalence of identified categories of differences and related subcategories in data among PSTs and ISTs group discussions:

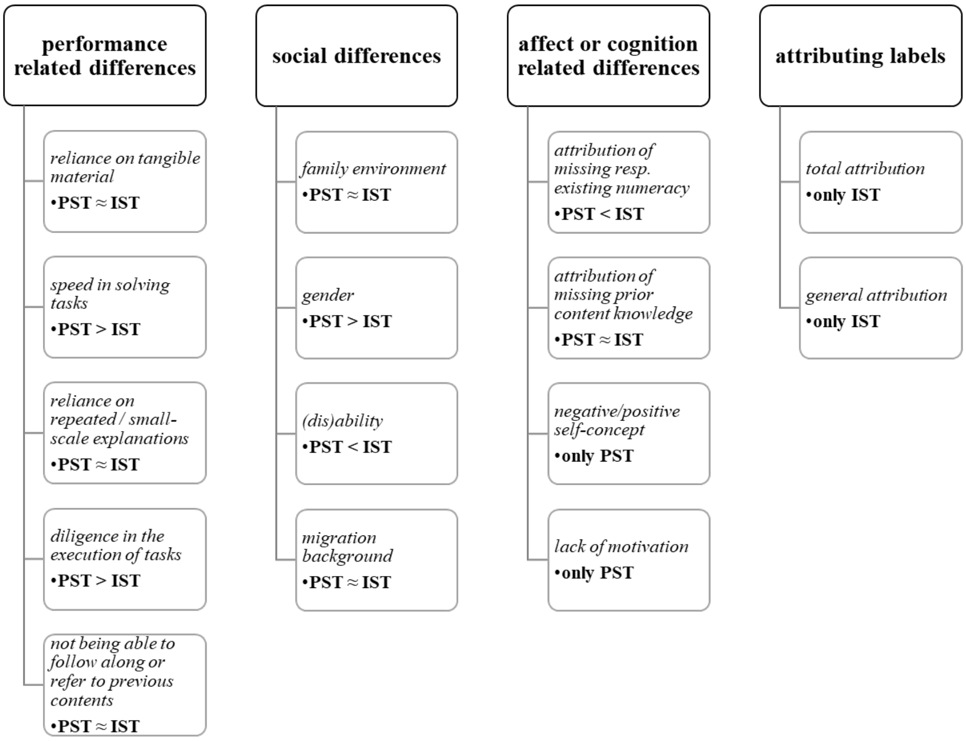



Figure 1: Prevalence in categories of differences with related subcategories.

**Preliminary results and interpretation**

The most noticeable similarity between PSTs and ISTs explicated differences is that they were constructed dichotomous without exception throughout all the analyzed data – e.g. learners were either characterized to show the distinguished performance or not, while the possibility of being in a process or state in between was not mentioned anywhere in the data. Within lines of social differences, no intersections were made and segments focusing on different aspects in learners' social identity are located separately in the data, not showing any connections between two or more aspects when referring to individuals or groups of learners. These findings might suggest that both PSTs and ISTs construct differences in learners as static or absolute without reflecting on the dynamic, evolving or intersectional state of learner's identity or ability. Especially within mathematics education, this stands in contrast to a perspective on learners as individuals in constructive learning processes. Another similarity across all data consists in the predominance of explicated differences that were analyzed as *'performance related'* and as *'social'* differences – in total, these two categories were in balance in between PSTs and ISTs data. The most common – and thereby possibly perceived as most relevant – explicated difference in learners across all data was *'reliance on tangible material'*, which was throughout aligned with "needing" hands-on-material to solve or understand tasks. This negative connotation reflects on a deficit-oriented view on learners without considering other possible applications of material such as argumentation tool while explaining own thoughts, helping others or as means of exploration. In a reversed perspective, not needing material marked capable mathematics learners and seemed to be considered as a desirable mode of operating in the classroom. Another commonly explicated difference in between learners was a *'reliance on repeated or small-scale explanations'* which both PSTs and ISTs used to distinguish those who understand a task or content immediately after the teachers' input and those who don't. This could contain a certain perspective on who is responsible for enabling understanding and allocating a lack of understanding not within teachers' instructions but rather into learner's ability – thereby immediate understanding might be framed as a desirable norm whereas needing further or other explanations is marked as 'other' without considering the necessity of adapting instructions based on inevitably varying starting points of learners.

Even though the predominance of *'performance related'* and *'social'* differences was shared among PSTs and ISTs, the distribution along associated subcategories as well as the prevalence in other categories provides diverging results when comparing PSTs and ISTs group discussions: Within *'performance related differences'* PSTs referred twice as much to students' speed in solving tasks and explicated more differences referring to learners' *'diligence in execution of tasks'*. This might suggest a stronger focus of PSTs on observable aspects of task execution (such as time, completeness, tidiness) instead of indicators of leaners' understanding during the process. Within the category of *'affect or cognition related differences'* only PSTs explicitly raised differences concerning a negative or positive *'self-concept'* and *'internalized attitudes towards math'*, which ISTs did not. This might suggest PSTs reference to common theories presented in their academic teaching which aim to explain interindividual differences during learning processes. *'Attributing labels'* to learners' identity as individuals or collectively as class were only used by ISTs, strikingly also along the dichotomy of "weak/slow" and "smart/quick" learners, highlighting them as peaks within the class and denying a middle field in between these two groupings. This is also remarkable since ISTs do not teach mathematics in the same groups of learners, but PSTs did. So, referring to individuals or groups of learners with such total or general attributions relied on the other ISTs in the group understanding who was meant without making it explicit. These attributions might be interpreted as means of collective communication, representing an established construction of differences in learners. Lastly, results concerning stated *'social differences'* showed diverging prevalence as well as varying references between the addressed social difference and mathematical learning. PSTs aligned differences within students' *'family environment'* to transmitted attitudes towards the subject and gained experiences involving mathematical activities during play and everyday life. ISTs framed differences within students *'family environment'* as relevant due to different levels of given support and repetition at home as well as performance pressure. These connotations entail different attributions towards the influence of family environment, resp. the given social and cultural background within students' homes. While PSTs seem to see students' family environment as initial condition affecting mathematical learning, ISTs seem to perceive it more as an accompanying and potentially supporting or aggravating aspect. Divergences between the explicated differences stemming from PSTs and ISTs were even more apparent within the subcategory *'gender'*: In only one discussion among ISTs, gender was referred to as a relevant difference for mathematics learning and only in context of performance, such as "boys don't do more than necessary". In PSTs' discussions on the other hand, gender was marked as relevant in all discussions and framed as a critical aspect in self-concept, as a part of mathematical aptitude and mathematical interest. Especially remarkable was the narrative of mothers' negative mathematical self-concept transferring to their daughters, which was independently presented in all PSTs' discussions. Even though this perception of genders' influence on mathematics learning may be subject of criticism, it is apparent that PSTs ascribe a higher relevance to connections between students' gender identity and their mathematical learning. In contrast, ISTs referred a lot more to diagnosed (learning) *'(dis)abilities'*, themselves including dyscalculia as a relevant difference in mathematics learners. Even though PSTs did also explicate this category of difference, it was much less apparent than in ISTs discussions. This might be due to PSTs lesser exposure to classroom practice and involvement with special needs students and therefore a lower connection with their perceived relevance – or it might be based in the thought of these learners being already marked or diagnosed as different which seemed so obvious that there was no need for explication. *'Migration background'* was the least mentioned subcategory within social lines of differences, its relevance mainly attributed to accompanying language barriers and once associated with missing everyday knowledge for problem solving. It is important to note, that the subcategory *'language'* was only explicated related to migration background and not referring to class. Furthermore, the forming of the subcategory *'(dis)ability'* including segments referring to dyscalculia is due to participants framing it as an aspect of special education needs. These two categories reflect the inductive category formation based on the analyzed texts without referring to a possible (other) theoretical outlining of these differences.

In conclusion, the highlighted results show shared constructed categories of differences, but varying ascriptions of relevance and their interweavement with mathematics learning between the discussions among PSTs and ISTs. PSTs focused a lot more on aspects of task execution and explicated more affect-related differences as well as highlighting gender as an important difference in mathematics learning. They seemed to refer a lot more to theories or concepts rooted in their academic teacher training (e.g. learning theories) for attributing meaning or relevance to the explicated differences. ISTs on the other hand seemed to focus more on experiences rooted in their practical experience, possibly relating differences more to expectations in mathematics learning such as using material or building on content knowledge. Additionally, they used more depictions of differences that can be interpreted as labels, such as "inclusive children" or total attributions without specifying the entailed meaning of these labels. These divergences between PSTs and ISTs might be due to their different bases of knowledge on which they (implicitly) refer to. While PSTs practical knowledge is mainly based in an academic and theoretical approach to teaching, ISTs practical knowledge is mainly dominated by everyday teaching experience at this time in their professional path. Subsequently, it appears to be crucial what the (implicitly) referred to practical knowledge contains of, revealing differences according to the extent of everyday teaching experiences building into practical knowledge. On the other hand, the influence of practical knowledge on marking differences as relevant is apparent throughout, regardless of the bases it is built upon.

### Limitations

The presented results allowed a deeper discussion of similarities and differences within emerging categories found in the group discussions of PSTs and ISTs and the lines of differences explicated therein. But results also call for a critical limitation and conclusion concerning further research steps. Firstly, all quantification of data is only applicable for the presented results and does only claim limited validity. The analysis of explicated differences is located in the sphere of PSTs and ISTs talking about their practical experience, reflecting a process of sense making and communication about practices and (implicit) norms. If and how the identified categories of differences influence classroom interaction needs to be analyzed in prospective studies. This also entails a critical assessment on the attribution of significance to 'social differences' in research's emphasis and the attribution of relevance made by teachers. Since the implicit meaning of explicated differences calls for more detailed empirical validation through reconstruction, it is not applicable to make a statement whether participants attribute the differences to learners' achievement in mathematics learning or to their presumed mathematical ability.

## CONNECTING 'DOING DIFFERENCE' TO 'SHARED SPACE OF EXPERIENCE'

Following the stated analytical frame of 'doing difference', all identified differences underly a logic of marking something as 'other' in contrast to 'normal/ expected' attributions in learners. While the normal or expected attributions in learners mostly remain implicit (since they are not mentioned explicitly), all explicated differences tend to consist of a specific comparative deficit or in marking them as special. Even though a reconstruction of implicit understandings of 'normal learners' cannot be provided in this paper, data discussion aimed at connecting the explicated differences as 'other' to possible counter horizons of implied 'normal'. In further steps, a reconstructive approach is taken to analyze negotiation processes accompanying the constructed differences to identify to which level individually proposed lines of differences are shared among the discussing group and to trace implicit levels of meaning within the discussion. In order to provide these insights, data of ISTs discussions will be additionally analyzed using Documentary Method, aiming at the reconstruction of participants' shared system of relevance within discussion. In line with the meta-theory of 'Praxeological Sociology of Knowledge', teachers tend to have similar practical experiences which provide orientation to action – this practical knowledge is based in a 'shared space of experience' and documents itself in joint narrations and depictions during discussion. Accordingly, a connection between constructed differences as products of the field-specific logic, according to the process of 'doing difference', and the constituting orientations leading this process, as documentation of a 'shared space of experience', can be provided. Tracing not only which differences in learners are constructed as relevant for mathematics education, but also how and why this is done, promises a deeper understanding of implicit norms held by mathematics teachers and proposes the possibility to make them accessible for self-reflective aspects in teacher professionalization.

## References

Bohnsack, R. (2010). Documentary Method and Group Discussions. In R. Bohnsack, N. Pfaff, W. Weller (Eds.), *Qualitative Analysis and Documentary Method* (pp. 99–124). Barbara Budrich.

Bräu, K. & Schlickum, C. (2015). *Soziale Konstruktionen in Schule und Unterricht*. Barbara Budrich.

Denessen, E., Honstra, L, Van den Bergh, L. & Bijlstra, G. (2022). Implicit measures of teachers' attitudes and stereotypes, and their effects on teacher practice and student outcomes: A review. *Learning and Instruction*, 78, Article 101437.

Gutstein, E. (2006). *Reading and writing the world with mathematics: Toward a pedagogy for social justice.* Routledge.

Kollosche, D., Marcone, R., Knigge, M., Penteado, M. & Skovsmose, O. (Eds.) (2019). *Inclusive mathematics education*. Springer

Hummel, A. & Reinhold, S. (2023). ‚Doing difference' in mathematics education. In *Proceeding of the 46th Conference of the International Group for the Psychology of Mathematics Education, Haifa, Israel,* Vol. 1, 362.

Hirschauer, S. (2014). Un/doing Differences. *Zeitschrift für Soziologie*, 43(3), 170–191.

Jacobs, V., Lamb, L. & Philipp, R. (2010). Professional noticing of children's mathematical thinking. *Journal for Research in Mathematics Education, 41* (2), 169–202.

Louie, N., Adiredja, A. & Jessup, N. (2021) Teacher noticing from a sociopolitical perspective. *ZDM Mathematics Education, 53*, 95–107.

Mayring, P. (2014). *Qualitative content analysis.* Beltz.

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

Rohn, D. (2013) Equity in Education: The Relationship between Race, Class, and Gender in Mathematics for Diverse Learners. *Urban Education Research and Policy*, 1, 12–22.

Walgenbach, K. (2014). *Heterogenität – Intersektionalität – Diversity in der Erziehungswissenschaft.* Barbara Budrich.

Voss, T., Kleickmann, T., Kunter, M. & Hachfeld, A. (2013). Mathematics teachers' beliefs. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Eds.), *Cognitive Activation in the Mathematics Classroom and Professional Competence of Teachers* (pp. 249 – 271). Springer.

# PROFILING INITIAL PEDAGOGICAL PRACTICES DURING MATHEMATICS PROFESSIONAL LEARNING AND DEVELOPMENT

Jodie Hunter[1], Julia L. Hill[2], Robyn Jorgensen[3], and Roberta Hunter[1]

[1]Massey University, New Zealand

[2]University of Melbourne, Australia

[3]University of Canberra, Australia

*Disparities in educational outcomes potentially indicate inequitable educational practices. This shows the importance of considering pedagogical practices in use in mathematics classrooms and the shifts in practice during professional learning. The study presented in this paper used a specially designed profiling tool to document the pedagogical practices used by 139 teachers in the first year of a professional learning initiative focused on ambitious teaching, culturally sustaining mathematics pedagogy, and mathematical wellbeing. The results showed that pedagogies focused on a supportive classroom environment were more evident than those related to ambitious or culturally sustaining mathematics pedagogy. We explain how profiling lessons can be used to identify areas that require more professional learning opportunities.*

## INTRODUCTION

Developing equitable outcomes for all learners in mathematics is an ongoing concern and challenge for educators, researchers, and policymakers (Hunter & Hunter, 2023; Louie, 2017). Persistent disparities in educational outcomes indicates widespread inequitable educational practices (Kennedy, 2016). In the context of New Zealand, both Māori and Pacific learners have long experienced structural inequities resulting in disparity in levels of mathematical achievement (Hunter & Hunter, 2023; May et al., 2019). Māori are indigenous to New Zealand, while Pacific people are closely related and include a heterogeneous grouping of recent arrivals from different island nations (e.g., Samoa, Tonga, Cook Islands, Niue, Tokelau, Fiji, Tuvalu) and multiple generations born in New Zealand. Disparity in mathematical achievement outcomes indicates a need to consider the pedagogical practices teachers are using in their mathematics classrooms and how to support teachers to change to pedagogical practices that better serve students from diverse cultural backgrounds.

Teacher professional learning and development (PLD) is commonly accepted as a key aspect of developing more equitable classroom contexts and in this way transforming educational systems (Guskey, 2002; Kennedy, 2016). However, supporting teachers to change pedagogical practices in mathematics teaching and learning is challenging with an additional difficulty related to how to both observe and document pedagogical practices and potential changes (Guskey, 2002; Shirrell et al., 2019). Differing theoretical frameworks of equitable and effective pedagogy are presented in research, however, in this study, we centre the need for teachers to develop their use of effective



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 89-96). PME.

pedagogical practices in mathematics through ambitious, culturally sustaining pedagogy that focuses on mathematical wellbeing (Hill et al., 2021; Lampert et al., 2010; Paris & Alim, 2014). Ambitious pedagogy is a form of inclusive teaching with key aspects including the use of cognitively demanding tasks coupled with opportunities for students to engage with mathematics disciplinary practices (Lampert et al., 2010; Youngs et al., 2022). Culturally sustaining mathematics pedagogy (CSMP) recognises that all learners bring strengths to the mathematics classroom from their social, cultural, and linguistic contexts. In this frame, educators provide equitable opportunities to learn mathematics by building on student strengths (Paris & Alim, 2014). Teaching for mathematical wellbeing involves pedagogy which aligns with student values and promotes positive feelings and functioning (Hill et al., 2021).

In New Zealand, a three-year PLD initiative called Developing Mathematical Inquiry and Communities (DMIC) (see following section) focuses on using mentoring and practice-based pedagogies to support teachers' uptake of pedagogical practices aligned with ambitious and culturally sustaining pedagogy that supports mathematical wellbeing. The study reported in this paper focuses on profiling the initial pedagogical practices used by teachers involved in the PLD through a structured observation tool modified from previous work on productive pedagogies (Lingard, Mills, & Hayes, 2003). Specifically, in this paper, we present a snapshot of the pedagogical practices used by teachers in the first year of the PLD to examine how such a tool can be used to both support the enactment of PLD and the development of tailored teacher support. The research question guiding the study presented in this paper is: (1) *What are the initial pedagogical practices used by teachers in a PLD intervention focused on ambitious mathematics teaching, CSMP, and mathematical wellbeing?*

## DMIC PROFESSIONAL LEARNING AND DEVELOPMENT INITIATIVE

Situated in the New Zealand context, DMIC PLD is a research-based professional development and pedagogical change initiative. This work has been funded by the New Zealand Ministry of Education and has grown and evolved in response to the persistent inequities for students from Māori and Pacific heritage in New Zealand. Those schools serving Pacific communities have been prioritised for inclusion in the PLD with many of these schools also having a significant proportion of Māori students. The initiative uses a whole-school approach and predominantly involves teachers of primary, middle, and lower secondary school students (Year One to Year Ten). All teachers involved in the PLD are provided with mathematical task resources which include sample teaching tasks, independent activities, and teacher notes outlining the key mathematical ideas and related mathematical language. Both learning activities outside of the classroom and within classroom mentoring during mathematics lessons are used in a complementary form. Outside of the classroom, professional development meetings involve teachers in exploring, discussing, and reflecting on pedagogical practices aligned with CSMP, ambitious pedagogy and mathematical wellbeing. For example, this may include activities such as reflecting and interrogating their own values and beliefs in order to consider pedagogical practices that align with the values of their students. Alternatively, teachers may be engaged in planning how to launch a cognitively challenging task or in anticipating student responses to tasks and rehearsing how to teach these while noticing student mathematical thinking and using practices such as making explanations or justifying reasoning. Within the classroom, dynamic in-the-moment mentoring is used with the teacher and mentor working together to co-construct mathematics lessons. Deliberate pauses are used both by the mentor and teacher during the lesson to enable professional conversations focused on reflecting and learning. The PLD is tailored over three years with a focus on schools beginning to independently sustain pedagogical practices through lesson study in the third year.

## DMIC PROFILING TOOL

To develop the profiling tool, we built on the work related to productive pedagogies in Queensland, Australia (Lingard et al., 2003; Sullivan et al., 2013). This initial research argued for a focus on pedagogical practices when considering quality educational and learning contexts given the influence these have on social and intellectual outcomes. We built on this work by modifying and expanding the productive pedagogies model to align with elements that are integral within the framework of DMIC. Three key dimensions were identified—intellectual quality, cultural connectedness, and supportive environment. The dimension of intellectual quality focuses on embedding high levels of mathematics content as well as mathematics disciplinary practices. Cultural connections refers to pedagogy that recognises students bring social, linguistic, and cultural knowledge and experience to the mathematics classroom which teachers can build upon as strengths. The final dimension of supportive environment refers to the creation of a learning environment where students feel a sense of belonging and are supported to participate in mathematics lessons. In each of the dimensions, we mapped aligned pedagogical practices as shown in Table 1 (see the findings section).

As noted in earlier research (Jorgensen et al., 2010; Sullivan et al., 2013) an important aspect of the tool is that it is a profiling instrument rather than a tool to assess teaching. An aim in this study was to document the initial pedagogical practices observed and enacted in the classroom to observe the presence (or absence) and strength of a pedagogical practice. A scale of one to five was used for each pedagogical practice with consideration of both the quality of the pedagogy and duration throughout the lesson. A score of one was used to indicate a total absence of the pedagogy in a lesson, through to five for the pedagogy being a strong feature of the entire lesson.

## METHODS

### Sample and data collection

We report on the observational data from 139 teachers during their first year of the DMIC PLD. Teachers were observed in either 2021 ($n$ = 61 teachers) or 2022 ($n$ = 78) on three separate occasions. The teachers were employed across 24 schools (23 primary/middle school, 1 secondary) situated in low to high socioeconomic neighbourhoods across New Zealand. The teachers had between 1 - 6 years total teaching experience. Aligned with the productive pedagogies method used in earlier studies (e.g., Lingard et al., 2003; Sullivan et al., 2013), each observation involved two observers separately observing the lesson in its entirety (typically between 45 minutes to 60 minutes). Prior to beginning the observations, all observers participated in a full day training workshop to learn how to use the profiling tool. Following the lesson, the observers independently scored the lesson for each pedagogy with nominal scoring that reflected the whole lesson. The observers then engaged in a discussion to moderate the scoring and agree on a joint score of between one to five for each pedagogy.

## RESULTS

Table 1 summarises the distributional data across the 3 pedagogical dimensions and corresponding practices of teachers during their first year of the DMIC PLD.

| Pedagogical dimension/ practice | Description | Mean (SD) | Min-Max |
|---|---|---|---|
| Intellectual quality dimension | | 2.42 (.72) | 1-4.3 |
| Challenging tasks | Challenging group-worthy teaching tasks to create opportunities for higher order thinking and/or to apply mathematics; tasks which are open-ended and/or have no immediately obvious solution. | 2.49 (.88) | 1-5 |
| Big ideas | Highlight key mathematical ideas and concepts both within tasks and in students' ideas and reasoning and support multiple connections across these. | 2.31 (.93) | 1-5 |
| Deep understanding | Support the development of deep understanding including both conceptual understandings and procedural fluency. Connect to general ideas of mathematics. | 2.08 (.92) | 1-4.5 |
| Substantive conversations | Productive communication (verbal/non-verbal) and sustained interactions between teacher to students and students to students focused on mathematical ideas and thinking. | 2.44 (.86) | 1-5 |
| Problematic knowledge | Focus on different perspectives in mathematics (equity/social justice) and viewing mathematical knowledge as a product of social, cultural, and political developments. | 1.91 (.93) | 1-4 |
| Language of maths | Accurate mathematical terminology and explicit attention to the meaning of mathematical vocabulary. | 2.91 (.81) | 1-5 |

| | | | |
|---|---|---|---|
| Mathematical practices | Mathematical disciplinary practices such as developing explanations, justifying, conjecturing, representing, generalising. | 2.65 (.72) | 1-5 |
| Teacher actions | Active facilitation of the lesson by anticipating, noticing, responding, and structuring student thinking and reasoning. | 2.4 (.88) | 1-4.7 |
| Lesson structure | Purposeful lesson structure with a launch, collaborative work and discussion, and connections to key mathematical ideas. | 2.4 (.98) | 1-5 |
| Cultural connectedness dimension | | 2.24 (.64) | 1.2-4.5 |
| Maths in culture | Contextual tasks connecting to everyday experiences and funds of knowledge with connections maintained throughout the lesson. | 2.43 (.89) | 1-5 |
| Problem based curriculum | Realistic contexts including numbers or situations which are plausible and would be solved using mathematics. | 2.66 (.91) | 1-5 |
| Trans-languaging | Active support for home language and English/Māori to facilitate communication. Language rich environment with different languages in mathematical repertoire. | 1.93 (.9) | 1-5 |
| Collectivism/ communalism | Explicit emphasis on working collectively to build ideas and active participation in learning. | 2.68 (.79) | 1-5 |
| Norms and values | Cultural values/norms acknowledged and embedded throughout the lesson with students participating in ways that maintain their cultural identity and integrity. | 1.49 (.83) | 1-5 |
| Supportive environment dimension | | 2.79 (.71) | 1-4.5 |
| High expectations | High expectations expressed with encouragement and affirmation of positive mathematical disposition towards intellectual challenge. Mistakes used as learning tools. | 2.43 (.84) | 1-5 |
| Engagement | Substantive engagement with mathematics, tasks, and peers. | 2.76 (.86) | 1-5 |
| Wellbeing | Calm, positive environment. Students have autonomy and their contributions are valued. | 3.07 (.97) | 1-5 |

| | | | |
|---|---|---|---|
| Grouping | Heterogeneous grouping with consideration of social needs and factors other than perceived ability. Status assigned throughout the lesson. | 3.03 (.78) | 1-5 |
| Social norms | Social norms are discussed and established. Pro-social behaviour evident in student interactions. | 2.71 (.9) | 1-5 |
| Inclusivity | Inclusion of all students despite differing learning or social needs, gender, culture, or language. | 2.77 (.83) | 1-4.5 |

Table 2: Description of the profiling tool with corresponding observational data

## DISCUSSION AND CONCLUSION

Across the pedagogical dimensions, lessons scored well in relation to implementing a supportive classroom environment ($M = 2.79$) and particularly student wellbeing ($M = 3.07$) and grouping ($M = 3.03$). Lessons scored less well on the ambitious pedagogy (intellectual quality dimension) ($M = 2.42$) with the use of problematic knowledge ($M = 1.91$) rated the lowest in this dimension. However, overall, the lowest rated pedagogical dimension was the use of culturally sustaining pedagogy (cultural connectedness dimension) ($M = 2.24$) with the lowest pedagogical practice the use of norms and values ($M = 1.49$). This is somewhat unsurprising as it aligns with early findings from the productive pedagogies framework (Ladwig, 1998; Lingard et al., 2001) with QSLRS studies reporting that teachers scored well on culture of care dimensions though less well on academic elements. Similarly, other studies using the Classroom Assessment Scoring System (CLASS) demonstrate teachers often score highest for fostering social emotional domains yet score lowest on producing high quality instructional support (La Paro et al., 2004; McGuire et al., 2016).

In the DMIC PLD initiative, a number of key pedagogies are emphasised so these profiles provide evidence for their relative uptake by teachers in the first year of the PLD. Rather than pointing to teachers' weaknesses, these signals areas that require more professional learning opportunities to develop impactful pedagogical practice both for the mentors working with the teachers and the teachers themselves. For example, in relation to the intellectual quality dimension, it is evident that with a mean score of 2.08 for deep understanding that teachers need stronger support and examples of how to achieve this throughout the lesson. This shows a need to engage teachers in developing their own understanding of mathematical concepts and key ideas embedded within tasks. In contrast, language of mathematics was scored at 2.91 which potentially shows that the mathematical task resources with the inclusion of related mathematical language were being well utilised by teachers.

In conclusion, research studies point to the importance of well-developed pedagogical practices for both teachers and students. For example, pedagogical competencies are linked to higher teacher wellbeing (Fernandes et al., 2019) and also enhanced students' engagement in mathematics (Pivot, 2023). Yet, there are few studies which have examined the shift in pedagogical practices longitudinally during PLD focused on ambitious and CSMP pedagogy and mathematical wellbeing. Future research will track teachers' pedagogical practices using this innovative tool over the length of the PLD and in the years that follow the PLD when a school is self-sustaining. This research is important considering high quality teaching practices are associated with higher student wellbeing (McCallum & Price, 2010) and achievement (Hattie, 2008). Thus, we conjecture teachers' pedagogical developments through DMIC will in turn promote students' mathematical wellbeing and achievement, aspects which we will also measure during the duration of the PLD initiative.

## References

Fernandes, L., Peixoto, F., Gouveia, M. J., Silva, J. C., & Wosnitza, M. (2019). Fostering teachers' resilience and wellbeing through professional learning: Effects from a training programme. *The Australian Educational Researcher, 46*(4), 681–698.

Guskey, T. R. (2002). Does it make a difference? Evaluating professional development. *Educational leadership*, *59*(6), 45-51.

Hill, J. L., Kern, M. L., Seah, W. T., & van Driel, J. (2021). Feeling good and functioning well in mathematics education: Exploring students' conceptions of mathematical wellbeing and values. *ECNU Review of Education*, *4*(2), 349–375.

Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.

Hunter, J., Hunter, R. (2023). Weaving together the threads of Indigenous knowledge and mathematics. *Educational Studies in Mathematics.*

Jorgensen, R., Grootenboer, P., Niesche, R., & Lerman, S. (2010). Challenges for teacher education: the mismatch between beliefs and practice in remote Indigenous contexts. *Asia-Pacific Journal of Teacher Education, 38*(2), 161-175.

Kennedy, M. M. (2016). How does professional development improve teaching? *Review of Educational Research*, *86*(4), 945-980.

La Paro, K. M., Pianta, R. C., & Stuhlman, M. (2004). The Classroom Assessment Scoring System: Findings from the prekindergarten year. *The Elementary School Journal*, 104, 409–426.

Ladwig, J. G. (1998). Authentic school reform. *Discourse: Studies in the Cultural Politics of Education, 19*(1), 113-119.

Lampert M., Beasley H., Ghousseini H., Kazemi E., Franke M. (2010). Using designed instructional activities to enable novices to manage ambitious mathematics teaching. In Stein M. K., Kucan L. (Eds.), *Instructional explanations in the disciplines* (pp. 129-141). New York, NY: Springer.

Lingard, B., Hayes, D., & Mills, M. (2003). Teachers and productive pedagogies: contextualising, conceptualising, utilising. *Pedagogy, Culture & Society, 11*(3), 399-424.

Lingard, B., Ladwig, J., Mills, M., & Bahr, M. (2001). *School Reform longitudinal study Vol1*. Brisbane: School of Education, The University of Queensland.

Louie, N. L. (2017). The culture of exclusion in mathematics education and its persistence in equity-oriented teaching. *Journal for Research in Mathematics Education*, *48*(5), 488-519.

May, S., Jang-Jones, A. & McGregor, A. (2019). *PISA 2018: New Zealand summary report - system performance and equity.* Ministry of Education.

McCallum, F., & Price, D. (2010). Well teachers, well students. *The Journal of Student Wellbeing*, *4*(1), 19-34.

McGuire, P., Kinzie, M., Thunder, K., & Berry, R. (2016) Methods of analysis and overall mathematics teaching quality in at-risk prekindergarten classroom. *Early Education and Development*, *27*(1), 89-109.

Paris, D., & Alim, H. S. (2014). What are we seeking to sustain through culturally sustaining pedagogy? A loving critique forward. *Harvard Educational Review, 84*(1), 85-100.

Pivot (2023). *Boosting student engagement in maths.* https://www.pivotpl.com/maths-report-2023

Shirrell, M., Hopkins, M., & Spillane, J. P. (2019) Educational infrastructure, professional learning, and changes in teachers' instructional practices and beliefs, *Professional Development in Education, 45*(4), 599-613.

Sullivan, P., Jorgensen, R., Boaler, J., & Lerman, S. (2013). Transposing reform pedagogy into new contexts: complex instruction in remote Australia. *Mathematics Education Research Journal, 25*(1), 173-184.

Youngs, P., Elreda, L. M., Anagnostopoulos, D., Cohen, J., Drake, C., & Konstantopoulos, S. (2022). The development of ambitious instruction: How beginning elementary teachers' preparation experiences are associated with their mathematics and English language arts instructional practices. *Teaching and Teacher Education*, *110*, 103576.

# MATHEMATICAL REASONING AND PROBLEM-SOLVING IN PISA 2022 – HOW DO PERFORMANCE PROFILES VARY ACROSS COUNTRIES?

Jenni Ingram, Gabriel Lee, and Jamie Stiff

University of Oxford, UK

*In PISA 2022, a new process subdomain was introduced focused on mathematical reasoning. This process was seen as the core of the problem-solving process that typifies PISA mathematics assessments. The results of PISA 2022 suggest that students in some countries have relative strengths specifically in mathematical reasoning, relative to the other problem-solving processes, while in other countries, this is an area of relative weakness. In this paper, we explore whether distinctive country profiles can be identified based on relative differences in performance on the four subdomain processes using Latent Profile Analysis. The profiles identified offer further support for considering the role of cultural and language contexts when comparing performance in international education studies.*

## INTRODUCTION

International large-scale assessment (ILSA) studies in education, such as PISA (Programme for International Student Assessment) and TIMSS (Trends in International Mathematics and Science Study), are highly influential drivers of policy and curriculum reform. In December 2023, the results of PISA 2022 were announced, which has been followed by many countries and education systems looking (selectively) towards the curricula, policies and practices of the higher-performing countries, using the PISA results as an external justification for reform (Johansson & Strietholt, 2019; Rojano & Solares-Rojas, 2018). For many of these higher-performing countries and education systems, these results are also being used as a justification for the success of recent reforms. Yet this process of turning to ILSAs for justification usually does "not draw on the detailed analytical insights that might be drawn from the PISA data" (Lingard, 2017, p.1). Throughout the remainder of this paper, the term country is used to describe all countries and education systems that participated in PISA 2022.

Around the world, mathematics curricula have become increasingly similar both in terms of the content areas included and the role of problem-solving and mathematical reasoning (Kadijevich et al., 2023; Valero, 2023). This homogenisation has been attributed to the growing globalisation of education policy and the influence of ILSAs, particularly ones that provide a 'ranking' of performance in mathematics and other key curriculum areas (Takayama, 2008). Of the two largest ILSAs focused on school mathematics, PISA assesses mathematical literacy with a focus on problem-solving in different task contexts, in contrast to TIMSS, where the focus is largely on mathematical content. The influence of both of these on mathematics curricula around



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 97-104). PME.

the world is visible in the increasing inclusion of particular mathematical content (e.g. statistics) and mathematical processes such as problem-solving as a major objective. That is, mathematics curricula around the world are including more and more of what is assessed in these ILSAs (Stacey et al, 2015).

In 2022 a new component, mathematical reasoning, was added to the mathematics framework for the PISA (OECD, 2023a). Described as being a core aspect of mathematical literacy, the PISA assessment framework "highlights the centrality of mathematical reasoning both to the problem-solving cycle and to mathematical literacy in general" (p. 23). This change aligns with recent research advocating that the main goal of a mathematics curriculum needs to be understanding which is reached through reasoning, with problem-solving as a means to develop this reasoning (Olivares et al., 2019). This new emphasis offers an opportunity to examine variations in students' strengths and weaknesses in the processes involved in problem-solving and mathematical reasoning in different country contexts.

One argument explaining the higher rankings of East Asian countries that has been seen in recent cycles of PISA is the tight focus of the national curricula in these countries on mathematics and science (and languages) rather than a more general and broader school curriculum, which are precisely the skills and content tested in TIMSS and PISA (Deng & Gopinathan, 2016). These arguments are often based on studies that show that students in these countries perform relatively low in measures of critical thinking and creativity (Lim, 2010) and also focus on the content of these curricula rather than the aims and objectives.

Problem-solving and reasoning, in contrast, are often considered in the aims, objectives and intended outcomes of a national curriculum. For example, Singapore was the highest-performing country in mathematics in PISA 2022, with an average score significantly higher than any other country (OECD, 2023b). The mathematics curriculum in Singapore has had problem-solving as the "primary aim" since 1990 (Fan & Zhu, 2007), with reasoning as part of the 'processes' component of the curricula. In Singapore, reasoning is the focus of the majority of the learning outcomes in the secondary mathematics curriculum (Serçe & Acar, 2021).

The results reporting the average subdomain scores published by the OECD show that some countries scored significantly higher or significantly lower in some subdomain areas than in others (https://www.oecd.org/pisa/). The research described above suggests that these relative strengths and weaknesses can be partly explained by differences in curricular focus across the different countries, particularly in relation to the mathematical content subdomains. The new subdomain of mathematical reasoning in PISA 2022 also offers an opportunity to examine these differences in relation to the mathematical processes.

This paper focuses on two research questions; What profiles in mathematics process performance in PISA 2022 can be identified? How do outcomes on the process subdomains differ across these profiles?

## METHODS

PISA uses a multi-stage adaptive test design meaning that participating students are required to only answer a subset of the items used. The mathematics and mathematics subdomains scores are calculated using plausible values for all students which have been scaled using multi-dimensional models. The mathematics scale and each of the process and content subscales are scaled together, meaning that analyses can consider relationships between the performances on each of the subdomain scores.

In this paper, we focus on the four process subdomains in mathematics; mathematical reasoning (or 'reasoning'), formulating situations mathematically (or 'formulate'), employing mathematical concepts, facts, and procedures (or 'employing') and interpreting, applying and evaluating mathematical outcomes (or 'interpreting'). Data were drawn from the 76 countries for which the subdomain scores are available. The average subdomain scores were calculated in a way that takes into account PISA's complex sample design, including the use of student weights and plausible values.

The subdomain scores correlate highly with the overall mathematics scale, with higher-preforming countries tending to perform highly on all scales, and lower-performing countries tending to perform lower. To address these general differences in performance, relative scores were obtained for each subdomain by subtracting the mean of the mathematical reasoning subdomain score from each of the process subdomain mean scores in each country. The relative scores therefore reflect the relative strengths or weaknesses in each subdomain to the mathematical reasoning mean score. Note, the overall mathematics score is not necessarily equal to the mean of the four process subdomain scores.

In addition, a homogeneity score was calculated for each country, which is the sum of the absolute values of the four subdomain scores minus the average of these four subdomain scores. A homogeneity score close to zero indicates that there was little variation between performance in the four subdomains, while a large value shows more pronounced strengths and weaknesses in particular areas. The average homogeneity score across all participating countries was 11.9 with a standard deviation of 5.0.

Latent profile analysis (LPA) was used to identify groups of countries with distinct performance profiles across the mathematics process subdomains. Countries are classified into groups based on membership probabilities estimated in the LPA model. The LPA variables were the four relative subdomain scores. Initially, the potential of three or four classes were examined based on a visual analysis of the relative subdomain scores, and the findings of a similar cross-country analysis of mathematics performance over time using TIMSS data (Johansson & Strietholt, 2019). Using the *mclust* package in R (Scrucca et al., 2023), the best fitting profile model was identified initially by comparing 3-group and 4-group models using the Bayesian Information Criterion and then using a Bootstrap Likelihood Ratio Test ($p$<0.001) to identify the optimal number of profiles. The final model was a 3-group spherical varying volume model.

## RESULTS

The Latent Profile Analysis identified 3 groups. The first group included 28 countries, while the second group included 27 countries, and the third group included 21 countries. The indicators for these 3 groups are shown in Table 1.

| Subdomain | Group 1 | Group 2 | Group 3 |
|---|---|---|---|
| Reasoning | 0.00 | 0.00 | 0.00 |
| Formulating | -3.03 | 3.95 | -5.15 |
| Employing | -0.11 | 7.05 | -4.08 |
| Interpreting | 3.78 | 4.69 | -6.07 |

Table 1: Indicators for the three group profiles for the subdomain processes.

Group 1 is characterised by a stronger mean score in interpreting and a weaker mean score in formulating, with reasoning and employing in between these two. The average homogeneity score for Group 1 was the smallest of the three groups (10.5), suggesting that these countries can also be characterised by relatively consistent scores across the four process subdomains, including the mathematical reasoning domain.

Group 2 is characterised by a weaker mean score in mathematical reasoning and a stronger employing mean score, with formulating and interpreting in between these two.

Group 3 is characterised by a relatively stronger mean score in mathematical reasoning in contrast to the other three process subdomains. Group 3 also had the largest average homogeneity score of 13.7 suggesting a wider spread of subdomain scores than in the other groups. However, the similar indicator values for the formulating, employing and interpreting subdomains suggest that the performance in these areas were similar to each other.

The most likely latent profile membership of the highest-performing countries is plotted in Figure 1. Countries have been ordered by their mean score for mathematics, and the figure shows the mean subdomain scores for each country with a mathematics mean performance significantly higher than the OECD average. The error bars represent the 95% confidence interval for the respective subdomain mean. (The graph showing the most likely profile membership of all participating countries will be shared in the presentation).


Figure 1: Subdomain mean scores for the highest-performing countries in PISA 2022.

The average mathematics score for countries most likely to be in Group 1 was 451, compared to an average score of 436 for Groups 2 and 3, suggesting that countries in Group 1 have stronger mathematics performance on average than countries in Groups 2 or 3. This was also the group with the highest average performance in mathematical reasoning with an average of 449, compared to an average of 430 for countries most likely to be in Group 2 and an average of 439 for countries most likely to be in Group 3.

The majority of the 6 top performing countries, which all performed significantly higher in mathematics than the other participating countries, were most likely to be in Group 2. However, the majority of the other countries most likely to be in Group 2 score significantly lower than the OECD average in mathematics (18 out of 27 countries). In contrast, less than half of the countries most likely to be in Group 1 had an average mathematics score significantly lower than the OECD average (13 out of 28 countries). English-speaking countries were also more likely to be in Group 1.

Over two-thirds of these higher-performing countries had homogeneity scores that were lower than the average homogeneity score across all participating countries. Of the top 6 performing countries, only Japan had a homogeneity score (13.1) that was larger than the average across all participating countries. Other higher-performing countries with an above-average homogeneity score included Sweden (21.0), New Zealand (15.8), Canada (13.9), Denmark (13.4), Austria (13.1) and The Netherlands (12.8).

## DISCUSSION

The analysis identified three profiles of performance in the PISA 2022 mathematics process subdomains. As well as illustrating different profiles of performance involving mathematical reasoning and problem-solving processes, they also illustrate some important cultural and linguistic differences. Johansson and Strietholt (2019) found that country-level strengths or weaknesses in mathematics content areas persist over time, suggesting that these profiles result from cultural contexts, including national curricula and national education policies and practices.

The majority of East Asian countries were most likely to be in Group 2, with a relatively stronger performance in employing and a relatively weaker performance in reasoning. The majority of European and English-speaking countries were most likely to be in Group 1, characterised by a stronger mean score in interpreting and a weaker mean score in formulating. These profiles suggest that culture and language have a substantial impact on students' mathematics performance in ILSAs. Furthermore, for countries most likely to be in Group 1, the mean interpreting score was often the highest score. This may reflect the emphasis on model validation and interpretation present in European mathematics education research (Niss, 1994; Geisler, 2021).

The highest 6 performing countries were all East Asian, which in the past has led to what Sellar and Linard (2013) referred to as the phenomenon of "looking East" to identify policies and practices explaining this high performance. The majority of these countries were most likely to be in Group 2, which was characterised by a weaker relative average performance in mathematical reasoning but a stronger relative average performance in employing; that is, employing mathematical concepts, facts and procedures. This suggests an emphasis on fluency with mathematical content, rather than on reasoning. However, these characteristics are not sufficient for high performance in mathematics more generally, as the majority of countries with this profile of performance scored significantly below the OECD average in mathematics.

The PISA 2022 mathematics framework puts mathematical reasoning at the centre of the problem-solving process. However, only the profile of countries in Group 1 had average performance scores in reasoning at the middle of the problem-solving process measures. The profile of Group 3 suggests a greater emphasis on mathematical reasoning than on the other problem-solving processes, while the profile of Group 2 suggests a greater emphasis on problem-solving processes than on mathematical reasoning. This analysis problematises the relationship between mathematical reasoning and the problem-solving processes and further research is needed to examine this relationship further.

The analysis in this paper assumes that PISA provides a valid measure of mathematics outcomes for different countries. While the participating countries influence the nature of the mathematics assessed, the content is based on what the OECD views as what students need to learn for today's (and tomorrow's) world (OECD, 2023a). This may align in different ways to country curricula. Furthermore, this analysis only examines the national picture, and it is also important to look within groups. The PISA 2022 results also found differences in mathematics performance between girls and boys, as well as depending upon a student's socioeconomic background. Finally, the differences in mean performance in each of the process subdomains were small in the majority of countries. While this paper has focused on comparisons with the performance in mathematical reasoning, the results also suggest that there is considerable variation in performance in the interpreting subdomain. Other models focusing on the relative performance in interpreting may lead to other performance profiles.

## References

Deng, Z. & Gopinathan, S. (2016). PISA and high-performing education systems: explaining Singapore's education success. *Comparative Education, 53(4),* 499–472. https://doi.org/10.1080/03050068.2016.1219535

Fan, L. & Zhu, Y. (2007). From convergence to divergence: the development of mathematical problem solving in research, curriculum, and classroom practice in Singapore. *ZDM, 39*, 491–501. https://doi.org/10.1007/s11858-007-0044-1

Geisler, S. (2021). Data-based modelling with experiments – students' experiences with model-validation. In M. Inpraitha, N. Changsri, & N. Boonsena (Eds.). *Proceedings of the 44th Conference of the International Group for the Psychology of Mathematics Education,* Vol 2, pp. 289–296. Khon Kaen, Thailand: PME.

Johansson, S. & Strietholt, R. (2019). Globalised student achievement? A longitudinal and cross-country analysis of convergence in mathematics performance. *Comparative Education, 55*(4), 536-556. https://doi.org/10.1080/03050068.2019.1657711

Kadijevich, D. M., Stephens, M., Solares-Rojas, A. & Guberman, R. (2023). Impacts of TIMSS and PISA on mathematics curriculum reforms. In Y. Shimizu & R. Vithal (Eds.), *Mathematics Curriculum Reforms around the World,* (pp. 359–374). Springer. https://doi.org/10/1007/978-3-031-13548-4_22

Lim, W. K. (2010). Asian Education Must Change to Promote Innovative Thinking. *Nature* 465 (7295): 157–157. https://doi.org/10.1038/465157a

Lingard, B. (2017). *The OECD's PISA: Purposes, impact and usage*. Paper presented at the International symposium "Educational governance in the era of international assessment". Japan Women's University. https://www.jeas.jp/jeas/wp-content/uploads/2017/10/Lingard-B.-The-OECD%E2%80%99s-PISA-Purposes-Impact-and-Usages.pdf

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 Academic Publishers.

OECD (2023a). *PISA 2022 Assessment and Analytical Framework,* PISA, OECD Publishing.

OECD (2023b). *PISA 2022 Results (Volume I): The State of Learning and Equity in Education*, PISA, OECD Publishing, https://doi.org/10.1787/53f23881-en

Olivares, D., Lupiáñez, J. L., & Segovia, I. (2021). Roles and characteristics of problem solving in the mathematics curriculum: a review. *International Journal of Mathematical Education in Science and Technology, 52*(7), 1079–1096.

Rojano, T., & Solares-Rojas, A. (2018). The mathematics curriculum design from an international perspective: Methodological elements from a comparative analysis. In Y. Shimizu & R. Vithal (Eds), *School mathematics curriculum reforms: Challenges, changes and opportunities. Proceedings of the twenty-fourth ICM1 Study conference* (pp. 475–482). International Commission on Mathematical Instruction.

Scrucca, L., Fraley, C., Murphy, T. B. & Raftery, A. E. (2023). *Model-based clustering, classification and density estimation using mclust in R*. Chapman and Hall. https://doi.org/10.1201/9781003277965

Serçe, F. & Acar, F. (2021). A comparative study of secondary mathematics curricula of Turkey, Estonia, Canada, and Singapore. *Journal of Pedagogical Research, 5*(1), pp. 216–242. https://doi.org/10.33902/JPR.2021167798

Stacey, K., Almuna, F., et al. (2015). PISA's influence on thought and action in mathematics education. In K. Stacey & R. Turner (Eds.), *Assessing mathematical literacy. The PISA experience* (pp. 275–306). Springer.

Takayama, K. (2008). The politics of international league tables: PISA in Japan's achievement crisis debate, *Comparative Education, 44*(4), 387–407. https://doi.org/10.1080/03050060802481413

Valero, P. (2023). Mathematical subjectivation: death sentence or chances for a terrestrial life. 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. 1, pp 53–68). University of Haifa, Israel: PME.

# ATTITUDES TOWARD MATHEMATICS AND GRAPHS INFLUENCE GRAPH REASONING AND SELECTION

Heather Lynn Johnson, Courtney Donovan, Robert Knurek, Kristin Whitmore, and Livvia Bechtold

University of Colorado Denver, USA

*We report on a mixed methods study in which we investigated college algebra students' attitudes toward mathematics and graphs in connection to their graph reasoning and graph selection. Students (n=599) completed a fully online survey of their attitudes toward math and graphs in conjunction with a fully online measure of their graph reasoning and selection for dynamic situations. Using structural equation modelling, we explored how students' attitudes might link to their graph reasoning and/or graph selection. We found that more positive attitudes toward mathematics and graphs linked to more quantitative forms of graph reasoning and more accuracy in graph selection.*

## INTRODUCTION

There is a complex relationship between students' attitudes toward mathematics and their mathematical thinking; it is essential that researchers engage in methods to embrace this complexity (Goldin et al., 2016). Drawing on DiMartino and Zan's (2010, 2011) model, we adopt a multidimensional view of attitudes toward mathematics, encompassing emotional disposition, perceived competence, and view of the subject. To theorize graph reasoning, we draw on the framework from Johnson et al. (2020), which puts forward four forms of reasoning: covariation, variation, motion, and iconic. To draw connections between students' attitudes and their graph reasoning and selection, we use structural equation modelling (SEM). SEM is a high-level statistical technique, in which researchers can demonstrate efficacy of theory-based models that relate different research-based constructs (Kline, 2023).

Our population comprises college algebra students (n=599), across three different U.S. postsecondary institutions. College algebra is a credit bearing course that often serves as a prerequisite to courses such as calculus, and functions and graphs are central to the course content. We investigate the following research question: To what extent does students' attitudes toward mathematics and graphs relate to the forms of their graph reasoning and/or the accuracy of their graph selection?

## THEORIZING STUDENTS' ATTITUDES TOWARD MATHEMATICS

DiMartino and Zan (2010, 2011) grounded their perspective on attitude in students' written narratives about their experiences, resulting in three interrelated dimensions: emotional disposition, perceived competence, and view of the subject. Emotional disposition referred to students' like, dislike, or indifference toward mathematics. Students' perceived competence referred to students' perceptions of their mathematical capabilities. Students' view of the subject referred to what mathematics meant for



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 105-112). PME.

students. Notably, Di Martino and Zan's perspective emerged from a goal to embrace complexities in students' attitudes, to push back against positive/negative dichotomies in investigations of students' attitudes.

## THEORIZING STUDENTS' GRAPH REASONING

The four-form graph reasoning framework from Johnson et al. (2020) distinguished between students' quantitative-based forms of graph reasoning (covariation, variation) and students' physical-based forms of graph reasoning (motion, iconic). The framework was developed to explain students' reasoning when interpreting and sketching graphs representing relationships between attributes in dynamic situations (e.g., a turning Ferris wheel). The covariation and variation constructs were rooted in Thompson's theory of quantitative reasoning (Thompson, 1994; Thompson & Carlson, 2017). In Thompson's theory, a quantity referred to a person's conception of some attribute as being possible to measure. For example, a person could separate the attribute of height from an object itself and conceive of how they might measure the height, even if they did not engage in any actual measuring. With covariation, Johnson et al. (2020) referred to students' reasoning about relationships between attributes, with at least a loose connection between their directions of change (e.g., height increases and decreases, while distance increases). With variation, they referred to students' reasoning about directions of change in a single attribute (e.g., height increases and decreases). With motion, they referred to students' reasoning about observable movements (e.g., Bell & Janvier, 1981; Kerslake, 1977) in the situation (e.g., a graph should show the path of the cart turning around the Ferris wheel). With iconic, they referred to students' reasoning about observable features (e.g., Clement, 1989; Leinhardt, 1990) in the situation (e.g., the Ferris wheel is curved, so my graph should be curved).

## METHODS

Our research design is a fully mixed, sequential, quantitative dominant status research design (Leech & Onwuegbuzie, 2009), with qualitative analysis preceding quantitative analysis. For data collection, we employed two fully online instruments, a survey of students' attitudes toward math and graphs (see Bechtold et al., 2022) and a measure of graph reasoning and selection for dynamic situations (MGSRDS) (Donovan et al., accepted; Johnson et al., in press). Both instruments were optimized for access on computers, tablets, and mobile phones. Students (n=599) completed the attitude survey and the MGSRDS concurrently, near the end of their college algebra course. Data collection occurred over three semesters (spring 21, fall 21, spring 22).

### Design of the attitude survey

To design our survey of students' attitudes toward mathematics and graphs (Table 1), we drew on Di Martino and Zan's (2010, 2011) conceptualization of attitudes toward math. This survey was an adaptation of a survey that Pepin (2011) administered, including three questions (Q1-Q3). Because we were investigating students' graph reasoning in conjunction with students' attitudes, we decided to also include items specifically connected to graphs (Q4, Q5). To allow students to express multifaceted responses, students were not forced to choose between like/dislike (emotional disposition) or can/cannot (perceived competence). Students responded to the survey with a text entry.

| Item |
| --- |
| Q1. I like/dislike mathematics because ____ |
| Q2. I can/cannot do mathematics because _____ |
| Q3. Mathematics is ____ |
| Q4. I like/dislike graphs because ____ |
| Q5. I can/cannot make sense of graphs because _____ |

Table 1: Survey of students' attitudes toward math and graphs

**Design of the MGSRDS**

The MGSRDS contains six items, with dynamic situation including a turning Ferris wheel, a person walking to a tree a back, a fishbowl filling with water, a cone growing and shrinking, a toy car moving along a square track, and two insects walking back and forth from home. Items appear in random order, and each item has two screens. On the first screen, there is a video animation of a dynamic situation (e.g., a turning Ferris wheel), written description of the attributes in the situation (e.g., In this situation, we will focus on the Ferris wheel cart's height from the ground and total distance travelled.), and a check for understanding. On the second screen, there are written instructions (e.g., Select the graph that best represents a relationship between the Ferris wheel cart's height from the ground and the distance travelled, for one revolution of the Ferris wheel.), and the video repeats. Then there are four graph choices representing relationships between attributes in the situation, and a text box for students to explain their graph choice. We have demonstrated validity for the MGSRDS (Donovan et al., accepted). For more on the design of MGSRDS items, see Johnson et al. (in press).

**Coding the attitude survey**

We used an interpretive approach to qualitative analysis to address complexities in students' attitudes toward mathematics and graphs. We coded students' attitudes along four categories: positive, mixed, negative, and detached (Table 2). The codes arose from our analysis of students' text responses (see Bechtold et al., 2022; Gardner et al., 2019). To code, we used a mix of machine and human coding. Our team hired a consultant to train a machine learning program based on our coding scheme. For responses receiving less than 70% confidence with machine coding, we brought in human coders (this tended to be about 30% of responses). With human coders, we used consensus coding (Olson et al., 2016); two people coded independently, then met to calibrate their codes, necessitating 100% agreement. After qualitative coding, we transformed the descriptive codes into numerical codes for statistical analysis. Larger values indicated more positive attitudes (0-detached, 1-negative, 2-mixed, 3-positive).

| Code | Description | Sample Response |
|---|---|---|
| Positive | Like/can | I can do mathematics because I've always been fluent in the 'language' of mathematics. |
| mixed | Combination of positive/negative | I am inbetween sometimes I do understand graphs and other times I can get very confused. |
| negative | Dislike/cannot | I dislike graphs because I tend to forget what the rules are to how functions and equations are placed. |
| detached | Separation from oneself | It's all just following the formulas step by step. |

Table 2: Attitude codes, descriptions, and sample responses

**Qualitative analysis: coding the MGSRDS**

We coded students' graph reasoning based on the four-form graph reasoning framework from Johnson et al. (2020): covariation (COV), variation (VAR), motion (MO), iconic (IC). To account for written responses that indicated limited evidence (LE) of reasoning, we added LE as a fifth code. Table 3 shows codes, descriptions, and sample responses. For the graph reasoning coding, we used only human coders. Again, we used consensus coding, which necessitated 100% intercoder agreement. After qualitative coding, we transformed the descriptive codes into numerical codes for statistical analysis. The values (0-LE, 1-IC, 2-MO, 3-VAR, 4-COV) indicated a hierarchy of graph reasoning (Donovan et al., accepted), with the largest values indicating quantitative graph reasoning (3-VAR, 4-COV).

| Code | Description | Sample Response |
|---|---|---|
| COV | relationships between directions of change in attributes | Total distance keeps increasing but the height increases then decreases |
| VAR | directions of change in a single attribute | The height increases, then decreases, and finally increases again |
| MO | Physical movement of objects in a situation | Shows the motion of the ferris wheel |
| IC | Physical features of a situation | If you connect the line, it becomes a circle just like the route it made |
| LE | Limited evidence | Just seems like the answer |

Table 3: Graph reasoning codes, descriptions, and sample responses

We coded students' graph selection using a spreadsheet. To guard against bias, we separated students' written explanations of their graph choice from their graph selections. The design of the MGSRDS included graph choices that were correct, partially correct, and incorrect. The partially correct graph choices accurately represented the direction of change in each attribute but did not accurately represent values of each attribute (for more, see Johnson et al., in press). Again, we transformed the descriptive codes into numerical codes for statistical analysis. Larger values indicated more positive attitudes (0=incorrect, 1=partially correct, 2=correct).

**Quantitative analysis: SEM**

SEM is a statistical technique that examines relationship patterns between variables that are modelled latently (Kline, 2023). Latent variables are preferred because they allow the variance between items to be examined instead of combining items into a mean score; An additional benefit of SEM is the ability to model multiple pathways with multiple dependent variables being tested simultaneously. To use SEM, researchers first develop a theory-based model. Then they determine whether data patterns fit their model. If there is unsatisfactory model fit, researchers modify and/or re-evaluate (Kline, 2023). Because of its complexity, SEM requires larger sample sizes than techniques such as multiple regression models.

The model for this study (Figure 1) includes two independent variables: attitudes towards mathematics and attitudes towards graphs. The independent variables predict two dependent variables, graph selection and graph reasoning. The four directional arrows in Figure 1 show this. To operationalize the constructs of attitudes toward mathematics and attitudes toward graphs, we used students' responses to questions 1, 2, 4, and 5 from the attitude survey (see Table 1). We did this because there was parallel structure in the design of the questions; one question about mathematics and graphs related to the dimensions of emotional disposition (Q1, Q4) and perceived competence (Q2, Q5), respectively. To operationalize the constructs of graph selection and graph reasoning, we used students' text responses explaining their graph reasoning and their graph choices for each of the six MGSRDS items.

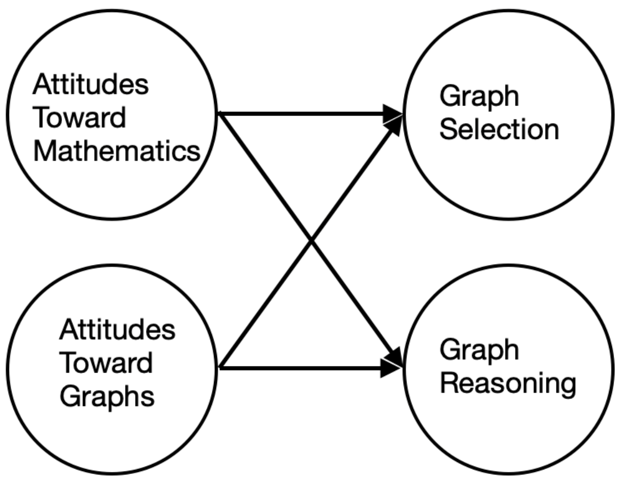



Figure 1: Conceptual model

We use three statistics to assess model fit: the chi-square goodness of fit, the Comparative Fit Index (CFI), the Root Mean Square Error of Approximation (RMSEA. The CFI addresses relative fit, assessing the model in comparison to a null, baseline model comprised of uncorrelated variables. CFI values of 0.90 and above provide sufficient evidence of good fit (Bentler & Bonett, 1980). The RMSEA addresses absolute fit, comparing a model is from an ideal. RMSEA values of 0.08 and below are considered acceptable fit (Browne & Cudeck, 1992). After assessing model fit, then we examine whether the items contributing to latent variables provide evidence of good fit (e.g., whether the six MGSRDS items contribute to the graph reasoning and graph selection constructs, and whether the four attitude survey items contribute to the attitude toward mathematics and attitudes toward graphs constructs). Standard regression weights of 0.30 or above are expected, with higher values indicating stronger contributions (Leech et al., 2014). If some items contribute at values lower than 0.30, they still may be included if they are significant and removing them does not improve the fit of the model.

## RESULTS

The conceptual model shown in Figure 1 demonstrated good fit, $\chi 2$ (99) = 154.93, $p < 0.001$, CFI = 0.96, RMSEA = 0.03. All items significantly contributed ($p < 0.01$) to the respective latent variable pathways. For graph reasoning, all six MGSRDS items contributed at values greater than 0.30 (values ranged from 0.58 to 0.77). For graph selection, four MGSRDS items contributed at values greater than 0.30 (values ranged from 0.33 to 0.48). The other two MGSRDS items contributed at values of 0.29 and 0.15. Removing these two items did not improve the model, thus we kept them. For attitudes toward mathematics and attitudes toward graphs, the four attitude survey questions (Q1, Q2, Q4, Q5, see Table 1) contributed at values greater than 0.30 (values ranged from 0.36 to 0.72). Hence, there was statistical support for our model.

All latent variable predictive pathways shown in the conceptual model (Figure 1) are significant. Like standardized regression weights, higher values indicate stronger relationships. Attitudes toward mathematics influences graph selection ($\beta = 0.80$, $p < 0.001$) and graph reasoning ($\beta = 0.64$, $p < 0.001$). Attitudes toward graphs influences graph selection ($\beta = 0.44$, $p = 0.006$) and graph reasoning ($\beta = 0.37$, $p = 0.002$).

## DISCUSSION

To begin, we asked: To what extent does students' attitudes toward mathematics and graphs relate to the forms of their graph reasoning and/or the accuracy of their graph selection? We found students' attitudes towards mathematics and attitudes toward graphs to influence their graph reasoning and graph selection. While all relationships were statistically significant, our results demonstrated that students' attitudes toward mathematics more strongly influenced their graph reasoning and graph selection than did students' attitudes toward graphs. Furthermore, the relationship between attitudes toward mathematics was stronger for graph selection than for graph reasoning. This also held for attitudes toward graphs. In all cases, more positive attitudes linked to more quantitative forms of graph reasoning and to more accuracy in graph selection. Furthermore, our results pointed to the interrelationships between the dimensions of emotional disposition and perceived competence within the constructs of attitudes toward mathematics and attitudes toward graphs, underscoring the complexity of the attitude construct posited by Di Martino and Zan (2010, 2011).

To contextualize our results, we discuss some limitations. First, we use students' written responses as proxies for their attitudes toward mathematics, their attitudes toward graphs, and their graph reasoning. Hence, there may be fuller aspects of these constructs not revealed by students' written responses. Second, while students completed the MGSRDS and attitude survey as part of their course, they may have viewed these instruments as "add-ons," and thus may have felt less investment in their responses (see also Johnson et al., in press). Third, our analysis conceptualizes attitudes as comprising only two of the dimensions of attitudes toward mathematics put forward by Di Martino and Zan (emotional disposition and perceived competence). Hence, our design simplifies the construct somewhat.

In conclusion, Goldin et al. (2016) suggested directions for future research to include the development of new instruments and the investigation of adults' attitudes toward mathematics. Our study furthered these research directions. In future studies, researchers could use the instruments we have developed with different populations.

**Acknowledgment.** This research was funded by the U.S. National Science foundation under DUE-2013186, Division of Undergraduate Education.

## References

Bechtold, L., Donovan, C., & Johnson, H. L. (2022). College algebra students' attitudes toward math and graphs: An exploratory factor analysis. In Lischka, A. E., Dyer, E. B., Jones, R. S., Lovett, J. N., 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* (pp. 1630-1634). MTSU. https://doi.org/10.51272/pmena.44.2022

Bell, A., & Janvier, C. (1981). The interpretation of graphs representing situations. *For the Learning of Mathematics*, *2*(1), 34–42. http://www.jstor.org/stable/40240746

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*(3), 588. https://doi.org/10.1037/0033-2909.88.3.588

Browne, M. W., & Cudeck, R. (1992). Alternative Ways of Assessing Model Fit. *Sociological Methods & Research*, *21*(2), 230–258. https://doi.org/10.1177/0049124192021002005

Clement, J. (1989). The concept of variation and misconceptions in cartesian graphing. *Focus on Learning Problems in Mathematics*, *11*(1-2), 77–87. https://eric.ed.gov/?id=EJ389508

Di Martino, P., & Zan, R. (2010). "Me and maths": Towards a definition of attitude grounded on students' narratives. *Journal of Mathematics Teacher Education*, *13*(1), 27–48. https://link.springer.com/article/10.1007/s10857-009-9134-z

Di Martino, P., & Zan, R. (2011). Attitude towards mathematics: a bridge between beliefs and emotions. *ZDM: The International Journal on Mathematics Education*, *43*(4), 471–482. https://doi.org/10.1007/s11858-011-0309-6

Donovan, C., Johnson, H. L., Knurek, R., Whitmore, K. A., & Bechtold, L. (accepted). Validating a measure of graph selection and graph reasoning for dynamic situations. *Journal of Mathematical Behavior.*

Gardner, A., Smith, A, & Johnson, H. L. (2019) Humanizing the coding of college algebra students' attitudes towards math. In Weinberg, A., Moore-Russo, D., Soto, H., & Wawro, M. (Eds.). *Proceedings of the 22nd Annual Conference on Research in Undergrad Math Ed* (pp. 1113-1114). OK. http://sigmaa.maa.org/rume/RUME22_Proceedings.pdf

Goldin, G. A., Hannula, M. S., Heyd-Metzuyanim, E., Jansen, A., Kaasila, R., Lutovac, S., Di Martino, P., Morselli, F., Middleton, J. A., Pantziara, M., & Zhang, Q. (2016). *Attitudes, beliefs, motivation and identity in mathematics education: An overview of the field and future directions*. Springer Nature.

Johnson, H. L., Donovan, C., Knurek, R., Whitmore, K. A., & Bechtold, L. (in press). Proposing and testing a model relating students' graph selection and graph reasoning for dynamic situations. *Educational Studies in Mathematics.*

Johnson, H. L., McClintock, E., & Gardner, A. (2020). Opportunities for reasoning: Digital task design to promote students' conceptions of graphs as representing relationships between quantities. *Digital Experiences in Mathematics Education*, *6*(3), 340–366. https://doi.org/10.1007/s40751-020-00061-9

Kerslake, D. (1977). The understanding of graphs. *Mathematics in School*, *6*(2), 22–25. http://www.jstor.org/stable/30212405

Kline, R. B. (2023). *Principles and Practice of Structural Equation Modeling*. Guilford.

Leech, N. L., Barrett, K. C., & Morgan, G. A. (2014). *IBM SPSS for intermediate statistics: Use and interpretation. (5th ed.)*. Routledge.

Leech, N. L., & Onwuegbuzie, A. J. (2009). A typology of mixed methods research designs. *Quality & Quantity*, *43*(2), 265–275. https://doi.org/10.1007/s11135-007-9105-3

Leinhardt, G., Zaslavsky, O., & Stein, M. K. (1990). Functions, graphs, and graphing: Tasks, learning, and teaching. *Review of Educational Research*, *60*(1), 1–64. https://doi.org/10.3102/00346543060001001

Olson, J., McAllister, C., Grinnell, L., Gehrke Walters, K., & Appunn, F. (2016). Applying constant comparative method with multiple investigators and inter-coder reliability. *The Qualitative Report*, *21*(1), 26–42. https://doi.org/10.46743/2160-3715/2016.2447

Thompson, P. W. (1994). The development of the concept of speed and its relationship to concepts of rate. In G. Harel & J. Confrey (Eds.), *The development of multiplicative reasoning in the learning of mathematics* (pp. 179–234). SUNY Press.

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.

# MODELLING PERFORMANCE USING FUNCTIONS: RELATION TO PERSON CHARACTERSITICS AND DIFFERENT SOLUTION APPROACHES

Michael Jonscher[1], Marielena Menzel[2], Sebastian Geisler[2], Stefanie Rach[1]

[1]Otto-von-Guericke University Magdeburg, [2]University of Potsdam, Germany

*Although mathematical modelling is undoubtedly a key competence, students often encounter challenges when working on modelling tasks. In a study with 122 tenth- and eleventh-grade students, we examined students' performance in modelling using functions by predicting it based on task values, self-concept, content knowledge, and prior achievement. In addition, we analysed students' solution approaches. Results indicate that students with high content knowledge and self-concept perform better in modelling. Both algebraic and graphical solution approaches enable precise solutions, but algebraic approaches are often abandoned. These results contribute the importance of both content knowledge and self-concept for modelling and indicate the potential of graphical assistance for algebraic solutions.*

## INTRODUCTION

Modelling is considered a mathematical key competence with great relevance for other scientific disciplines, everyday life, and society (Niss, 1994). While modelling is undeniably important, as illustrated by its long tradition in the PME (e.g., Kaiser & Schukajlow, 2022), it poses a challenge for students (Blum & Leiss, 2007). Especially in times of climate change and pandemics, it seems crucial to promote students' competence in modelling with functions.

Students reported the lowest motivation for modelling problems compared to dressed up word problems and intra-mathematical problems (Krawitz & Schukajlow, 2018). As students with higher motivation probably achieve higher performance in solving problems (see Heinze et al., 2005), it seems crucial to further investigate the role of person characteristics when modelling using functions. As modelling can be conceptualised by using mathematical concepts to describe real-world situations, it is plausible that content knowledge is important in modelling processes (see also Holenstein et al., 2022).

Our goals in this study are to examine the impact of person characteristics on students' performance and detect solution approaches when modelling using functions. As person characteristics, we focus on task values, mathematical self-concept, content knowledge, and prior achievement in mathematics. As little is known about students' solution approaches when modelling using functions, this study aims to explore which approaches contribute to successful modelling.



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 113-120). PME.

## THEORETICAL BACKGROUND

Mathematical modelling describes the process of solving real-world problems by using mathematics. Referring to Blum and Leiss (2007), an idealized modelling process can be described as a cyclic process that begins and ends in reality and consists of seven steps. First, the real-world problem needs to be understood, then simplified, and structured. After building a mathematical model, the problem can be solved using mathematical procedures. The obtained mathematical solution must be interpreted in order to derive a real result. After validating the real result and the whole modelling process as suitable, the solution for the real-world problem can be presented.

### Person characteristics and performance in modelling

Schukajlow et al. (2022) report about initial studies that show certain affective constructs, e. g., task values and self-beliefs, to be related to performance in modelling. In the situated expectancy-value theory by Eccles & Wigfield (2020), task values are defined as an important predictor of a persons' performance consisting of four components: attainment value, reflecting the personal importance and proficiency of the task; intrinsic value, signifying personal interest and anticipated enjoyment; utility value, covering the task's relevance in daily life, career, and other aspects; and costs, reflecting the tasks' negative aspects. In research, it is discussed if costs belong to task values or not (Muenks et al., 2023). Self-beliefs such as self-concept or self-efficacy expectations refers to an individual's beliefs, perceptions, and attitudes about their own abilities, competences, and identity in a domain (Eccles & Wigfield, 2020) and relate to modelling performance (Holenstein et al., 2022). In this contribution, we focus on self-concept and task values, which are considered as rather stable constructs (e.g., Gaspard et al., 2015).

In addition, we take into account content knowledge and prior achievement for explaining differences in students' performances. Under the term content knowledge, we subsume conceptual and procedural knowledge about functions, which is assumed to play a major role in using functions for real-world problems (Siller et al., 2022).

### Solution approaches when modelling

Kraemer et al. (2012) showed that students, who possess multiple solution methods, are more successful in handling modelling tasks. They identified five solution approaches: *algebraic*, i.e. setting up and solving equations; *graphical*, i.e. graphing functions in a coordinate system and interpreting the results; *exemplary*, i.e. inserting specific values and calculating accurately to answer appropriately; *numerical*, i.e. creating a value table focusing on relevant domains; and *content* approaches, i.e. employing specific real-world terms and their relations. While approaches differ in the expected precision of their solutions (Ainsworth, 1999), they also have varying levels of difficulty, where especially algebraic solutions appear to be extremely challenging for students (Galbraith & Stillman, 2006).

Kraemer et al. (2012) provided further insights into students utilizing these approaches: even though algebraic approaches allow for highly precise solutions, it is observed that students rarely utilize them. Whereas exemplary approaches seem suitable for introducing students to mathematical modelling with linear functions, graphical approaches were chosen extremely rarely by students. As content approaches were commonly employed by students and consistently allowed to find better solutions, it is recommended to particularly introduce this approach to students. Considering the high potential of precise solutions offered by algebraic approaches and students' difficulties with them, it seems important to further explore their conclusion that exemplary and numerical solutions, in combination with graphical illustrations, could serve as a promising foundation for transitioning to algebraic approaches in modelling.

## THE CURRENT STUDY

As part of the project Ex2MoMa – Experiments to foster Modelling Competences and Motivation in Mathematics – this study was designed to examine the impact of person characteristics on students' performance and to uncover solution approaches when modelling using functions.

### Research Questions

1. How do students' content knowledge, prior achievement, mathematical self-concept, and task values predict the performance in a modelling task?
2. Which solution approaches do typically lead to successful and unsuccessful modelling processes?

## METHODS

### Sample and design

Our sample comprises of 122 students from nine grammar schools (grades 10 and 11, *M*(age)=16.11, 65% girls). First, we measured students' prior knowledge concerning linear and exponential functions and the students worked on modelling task 1, involving linear functions. After one week, students reported their values in mathematics, their mathematical self-concept, and their last grade as an indicator of prior achievement in mathematics; in addition, they worked on modelling task 2, involving exponential functions. The students were familiar with typical characteristics of linear and exponential functions.

### Instruments

*Modelling tasks*: Task 1 (see figure 1) deals with the consumption of the battery of a car which can be modelled by a linear function, Task 2 with the sleeping quality on the Mont Blanc which can be modelled by an exponential function. To rate students' performance in modelling tasks, Schukajlow et al. (2023) developed an extensive coding scheme. This coding scheme was utilized for rating geometrical modelling tasks but appears to be well adaptable for any mathematical contexts. Based on the modelling cycle as described above, a total score for correctness ranging from 0 to 4 is possible, with one point allocated for each of the specified steps: identifying a suitable mathematical model, setting up the mathematical model, mathematical calculations and interpretation. Based on this scheme, we developed an extensive coding scheme for these tasks. Up to four points could be reached for every task solution. The interrater reliability for both modelling tasks is good: task 1 cohen's κ = .84, task 2 cohen's κ = .80. The descriptive statistic for students' solutions (task 1: $M = 0.71$ ($SD = 1.05$) and task 2: $M = 1.89$ ($SD = 1.44$) shows a floor effect for task 1.

**Battery Consumption**

For an electric car, the table below shows the current battery consumption based on the distance travelled. What distance can the car still cover before it needs to be recharged? Describe your solution approach.

| | | | | | | |
|---|---|---|---|---|---|---|
| Distance travelled in km | 0 | 50 | 100 | 200 | 300 | 350 |
| Battery Consumption in kWh after distance travelled | - | 8.2 | 15.8 | 33.7 | 49.0 | 57.7 |
| Remaining Battery in kWh after distance travelled | 96.9 | 88.7 | 81.1 | 63.2 | 47.9 | 39.2 |

Figure 1: Modelling task 1 (translated in English).

*Person characteristics*: Table 1 shows an overview about the instruments to measure person characteristics in order to predict the performance in the modelling tasks. To measure content knowledge, we developed a test, consisting of seven items that address the characteristics of linear and exponential functions. In addition, students were asked to report their last grade in mathematics (from 1 deficient to 5 very good) and rate statements concerning their mathematical self-concept (Arens et al., 2011) and task values (e.g., Gaspard et al., 2015) on a six-point likert scale (*1=totally disagree*, *6=totally agree*). The combined scale task values consists of the facets intrinsic value, attainment value, and utility value. No floor or ceiling effects were found concerning the person characteristics and acceptable to high internal consistency. The person characteristics correlate on a medium to high level.

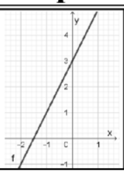
| Construct | Item example | Number of items | Cronbach's α |
|---|---|---|---|
| Content knowledge | Given is the graph of function f. Mark the equation that fits to the graph. ☐ $y = -2x + 3$ ☐ $y = 2x + 3$ ☐ $y = 3x + 2$ ☐ $y = 3x - 2$  | 7 | .75 |
| Prior achievement in mathematics | *(Last grade in mathematics)* | 1 | - |
| Mathematical self-concept | In mathematics, I am good. | 4 | .95 |
| Task values | I find mathematics exciting. | 19 | .95 |

Table 1: Internal consistency and further insights in person characteristics.

## Data Analysis

As performance in modelling task 1 varied only slightly between students, we used students' performance on task 2 for answering research question 1. We conducted linear regression analyses in SPSS, version 29, with performance as the dependant variable and person characteristics as independent variables. To answer research question 2, students' solutions concerning task 1 were used, as these solutions provide deep insights into why students were not successful in solving this task. Besides, successful solution approaches provide ideas how students can be supported in class.

## RESULTS

Linear regression analyses indicate that both content knowledge and prior achievement in mathematics predict the performance in a modelling task (table 2, model 1). Moreover, mathematical self-concept also predicts performance whereas task values do not explain variances in performance when controlling self-concept (model 2). Integrating the substantial predictors of performance in one analysis, both content knowledge and self-concept remain as significant contributors to performance.

| | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| Content knowledge | .39*** | - | .29** |
| Prior achievement in mathematics | .20* | - | .02 |
| Mathematical self-concept | - | .46*** | .36** |
| Task values | - | .02 | - |
| $R^2$ | .26 | .23 | .31 |

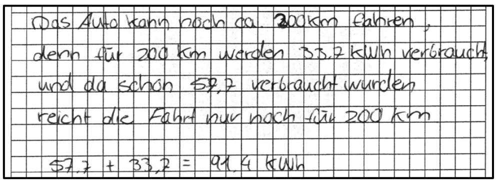
Table 2: Standardized regression parameters of linear regression analyses, performance of task 2 as dependant variable; *** $p < .001$, ** $p < .01$, * $p < .05$.

To explain why these person characteristics coincide with the performance in a modelling tasks and identify typical students' solution approaches, we screened all solutions, in particular concerning task 1. Because of page restrictions, we focus three different approaches:

*1. Algebraic solution:* Students developed a mathematical model by formulating the equation of a linear function and determining its parameters through calculations. To calculate the remaining range of the car, this linear function was set to zero and solved. Generally, the algebraic approach enables very precise solutions. Only a very few students were able to successfully utilize this approach, those who did, demonstrated a high modelling performance. However, uncommonly this approach was initiated by students but frequently abandoned in favour of using a content approach.

*2. Content solution*: Students approximated a solution graphically based on the provided values out of the given table. When using this approach, typically a linear relationship between battery consumption and the distance travelled was implicitly assumed. Figure 2 displays a student's attempt to add two values in a manner that achieves the best possible match with the initial remaining battery capacity. This approach was by far the most recent used approach. Students who utilized it, often only showed a low level of modelling performance.

*3. Graphical solution*: Students drew a coordinate system with the distance travelled on the x-axis and the remaining battery on the y-axis. They plotted the value pairs from the given table and, by visual estimation, drew a straight line that matched the points best (see figure 3). By identifying the x-intercept, a solution for the total range of the car was determined. When utilizing the graphical approach, usually students demonstrated a high level of modelling performance. Similar to algebraic solutions, this approach allowed very precise solutions. Only a few students used it but in contrast to the algebraic approach, the graphical approach was rarely abandoned.




"The car can drive another 200 km, because 33.7 kWh are consumed for 200 km. As 57.7 [kWh] already have been consumed, the journey is only sufficient for 200 km.

57.7 + 33.7 = 91.4 kWh".

Figure 2: A students' content solution.

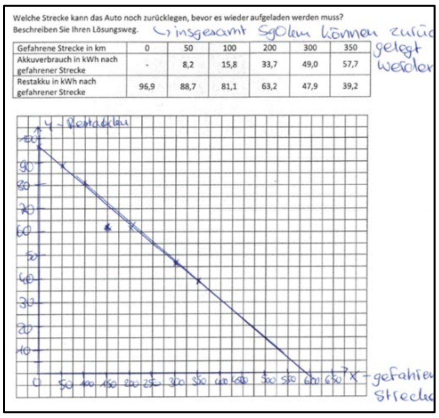



Figure 3: A students' graphical solution: "A total of 590 km can be covered."

## DISCUSSION

The starting point of our study was the phenomenon that using mathematics for describing real-world problems is an important activity but that students report substantial problems in solving modelling tasks (Blum & Leiss, 2007). In this study, we concentrate on modelling a certain situation with linear and exponential functions. The results show that students with higher content knowledge and a higher mathematical self-concept demonstrate better performance in solving the modelling task. The prediction based on the prior achievement in mathematics is nearly fully explained by mathematical self-concept and content knowledge. Task values, on the other hand, explain little to no additional variances in performance.

In-depth analyses of the students' solutions provide insights what problem solving strategies could be provided to students who struggle when searching a suitable solution. Many students using the content solution approach were unsuccessful due to insufficient precision in their results. Kraemer et al. (2012) propose that using content solutions is an extremely valuable strategy for mathematical modelling. However, for students it can be challenging to recognize in which cases this kind of solutions are sufficient. While the content approach often does not allow precise solutions, at least it still can be important for checking the conducted modelling process. Therefore, it seems crucial to communicate in class how content solutions should be utilized in terms of modelling and to derive normative rules for solving modelling problems.

Both algebraic and graphical solution approaches enable suitable modelling processes and exact solutions. In order to foster students' modelling competences, it is advisable to provide both of them. Although graphical solutions often allowed suitable modelling processes that rarely were abandoned prematurely (see also Galbraith & Stillman, 2006), algebraic solutions may be more advantageous in different contexts, resulting in more appropriate modelling. The close relationship between content knowledge and performance supports the explanation that only students with high knowledge concerning characteristics of functions can successfully apply the algebraic approach (e.g., Kraemer et al., 2012). Content knowledge serves as a resource in the modelling process. Furthermore developing a mathematical model by formulating a function appeared extremely challenging for students, when utilizing algebraic approaches. The close relationship between mathematical self-concept and performance in the modelling task suggests that students with higher self-concept are more confident to succeed these challenges instead of giving up with less optimal solutions, underpinning Holensteins' (2022) findings. In this step, students could benefit from graphical assistance, because it offers additional cues for determining parameters and appears to be more accessible for students. This emphasizes the assumption of Kraemer et al. (2012) that graphical assistance appears to have high potential in supporting students to successfully model with algebraic solution approaches.

Our results are limited by the fact that students' performance in task 1 differ only little so that only the performance in task 2 can be used to identify predictors. Task values are conceptualised as a global measure and in further studies, the single facets of task values may explain why some students perform better when solving modelling tasks than other students. As we concentrate on the product of the modelling process, we can't explain in which way students use content knowledge in the modelling process.

In sum, this study provides insights into students' performance and approaches when modelling real-world situations using functions. These insights can be used to support students in their modelling process.

**Funding:** The project is funded by the German federal ministry of education and research.

## References

Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education 33*, 131–152.

Arens, K., Trautwein, U., & Hasselhorn, M. (2011). Erfassung des Selbstkonzepts im mittleren Kindesalter: Validierung einer deutschen Version des SDQ I1 [Self-concept measurement with preadolescents: Validation of a German version of the SDQ I]. *Zeitschrift für Pädagogische Psychologie, 25*(2), 131–144.

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.

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*(4), 101859.

Galbraith, P. L., & Stillman, G. (2006). A framework for identifying student blockages during transitions in the modelling process. *ZDM Mathematics Education, 38*, 143–162.

Gaspard, H., Dicke, A.-L., Flunger, B., Schreier, B., Häfner, I., Trautwein, U., & Nagengast, B. (2015). More value through greater differentiation: Gender differences in value beliefs about math. *Journal of Educational Psychology*, *107*(3), 663–677.

Heinze, A., Reiss, K., & Rudolph, F. (2005). Mathematics achievement and interest in mathematics from a differential perspective. *ZDM Mathematics Education*, *37*(3), 212–220.

Holenstein, M., Bruckmaier, G., & Grob, A. (2022). How do self-efficacy and self-concept impact mathematical achievement? The case of mathematical modelling. *British Journal of Educational Psychology, 92*(1), 155–174.

Kaiser, G., & Schukajlow, S. (2022). Innovative Perspectives in Research in Mathematical Modelling Education. In C. Fernández, S. Llinares, A. Gutiérrez, & N. Planas (Eds.), *Proc. 45th Conf. of the Int. Group for Psychology of Mathematics Education* (Vol. 1, pp. 147-176). PME.

Kraemer, J., Schukajlow, S., & Blum, W. (2012). Bearbeitungsmuster von Schülern bei der Lösung von Modellierungsaufgaben zum Inhaltsbereich Lineare Funktionen [Students' processing patterns when solving modelling tasks in the content area of linear functions]. *mathematica didactica, 35*, 50-72.

Krawitz, J., & Schukajlow, S. (2018). Do students value modelling problems, and are they confident they can solve such problems? Value and self-efficacy for modelling, word, and intra-mathematical problems. *ZDM Mathematics Education, 50*, 143–157.

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 Academic Publishers.

Muenks, K., Miller, J. E., Schuetze, B. A., & Whittaker, T. A. (2023). Is cost separate from or part of subjective task value? An empirical examination of expectancy-value versus expectancy-value-cost perspectives. *Contemporary Educational Psychology, 72*, 102149.

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 Psychology of Mathematics Education* (Vol. 1, pp. 158-162). PME.

Schukajlow, S., Krawitz, J., Kanefke, J., Blum, W., & Rakoczy, K. (2023). Open modelling problems: cognitive barriers and instructional prompts. *Educational Studies in Mathematics, 114* (3), 417–438.

Siller, H.-S., Elschenbroich, H.-J., Greefrath, G., & Vorhölter, K. (2022). Mathematical modelling of exponential growth as a rich learning environment for mathematics classrooms. *ZDM Mathematics Education, 55*, 17–33.

# A PRELIMINARY SYSTEMATIC REVIEW ON HOW PRODUCTIVE STRUGGLE IS DEFINED IN MATHEMATICS EDUCATION RESEARCH

Nitchada Kamlue and Laura R. Van Zoest

Western Michigan University

*This systematic review investigated how* productive struggle *was defined in studies investigating productive struggle in mathematics learning. Following PRISMA guidelines, we identified 10 such peer-reviewed journal articles from the Scopus database from 2007 to 2023. We reported (a) (proxy) definitions of productive struggle for each study; (b) structural elements across the definitions—subject, action, object, and aim; and (c) synthesizing aspects across the definitions—definition foci and features of the objects. Finally, we initiated the process of rethinking together how to investigate what it means for mathematics learners to engage in productive struggle by sharpening the productive struggle construct.*

## INTRODUCTION

*Productive struggle* has become a popular phrase in mathematics education in the United States. The first use of "productive struggle" in the Scopus database related to mathematics was by Warshauer in 2015 and by the end of 2023 it was mentioned in 274 documents. This increasing use may be in part because "support[ing] productive struggle in learning mathematics" is one of the US National Council of Teachers of Mathematics' eight mathematical teaching practices ([NCTM], 2014, p .10). NCTM (2014) cited three research documents to support their claims about productive struggle: Hiebert and Grouws' (2007) handbook chapter on teaching and learning, Kapur's (2010) study of productive failure, and Warshauer's (2011) dissertation on productive struggle. Warshauer (2011) is the only of these three documents that reports on a study about productive struggle. Kapur (2010) made no mention of *productive struggle* (although the construct of *productive failure* is closely related) and Hiebert and Grouws (2007) identified *struggle with important mathematics* as a feature of teaching that research has found to support students' conceptual understanding. Hiebert and Grouws explicitly defined *struggle* to mean "students expend effort to make sense of mathematics, to figure something out that is not immediately apparent" (p. 387). Warshauer (2011) quoted Hiebert and Grouws (2007)'s definition of *struggle* when defining her use of *productive struggle*.

Kaiser and Schukajlow (2023) called for more literature reviews in mathematics education. Given the need for researchers to investigate *productive struggle* and the lack of an established definition of what is meant by the term, this seemed a useful time to complete a systematic literature review on the definitions of productive struggle currently in use. Thus we report here on a systematic literature review of the question,



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 121-128). PME.

*How is productive struggle defined in studies investigating productive struggle in mathematics learning*?

# METHODS

Our study is a systematic literature review—a study that has "systematic and rigorous search procedures" (Kaiser & Schukajlow, 2023, p. 2). We drew on recently published recommendations for completing literature reviews (e.g., Kaiser & Schukajlaw, 2023) and followed the PRISMA (2020) checklist and flowchart.

We had two major inclusion and exclusion criteria for our review: (a) we were interested in research studies that investigated productive struggle in mathematics learning, rather than simply mentioning the term; and (b) we wanted to focus on rigorous studies. As have other mathematics education researchers (e.g., Nieminen et al., 2023; Phan et al., 2022), we started with the Scopus database because it "is the largest abstract and citation database of peer-reviewed literature," covering over 25,000 journals across all disciplines (Scopus blog, 2023, homepage). After our initial comparison with the ERIC and Education Source databases revealed no additional peer-reviewed journal articles that met our criteria, we limited our search to Scopus for this preliminary systematic review. To meet our first criterion, we searched the article title, abstract, and keywords for any mention of "productive struggle" or mathematics (using "math*" to capture variations of the term). To meet our second criterion, we used publication in peer-reviewed journals as a proxy for "rigorous." Because Hiebert and Grouws' (2007) article is often credited with drawing the field's attention to the idea of struggle in mathematics education, we used 2007 as the starting point for our search. We included articles that had a publication date through the end of 2023. Figure 1 illustrates our process of identifying studies and the numbers that resulted at each step. Our initial search identified 30 articles. We first screened these 30 articles by reading their title, abstract, and keywords carefully to decide whether the research investigated productive struggle in mathematics learning. We excluded a total of 18 articles: (a) two because they were suggestions for practitioners rather than reports of research studies; (b) three because they focused solely on investigating beliefs or attitudes; and (c) thirteen because they did not investigate productive struggle in mathematics learning (e.g., investigating teachers' productive struggle when learning to design a mathematics curriculum, rather than when learning mathematics). We then considered the full text of the 12 remaining studies to verify that they did investigate productive struggle in mathematics learning. Two studies did not meet this criterion and thus were excluded from our data set. The remaining 10 studies were included in our review.

Our analysis included: (a) identifying a definition of productive struggle used in each study; (b) deconstructing the structural elements of each definition of productive struggle; and (c) synthesizing across those structural elements. For (a), we first searched each article for an explicit mention of the definition of productive struggle used in the study. We considered it an *explicit definition* if there was a clearly stated definition and it was clearly stated that it was the definition used in the study. If a definition was clearly stated, but it was not clearly stated that it was the definition used in the study, we searched the article for confirming and disconfirming evidence of its use. If there was no disconfirming evidence, we identified the definition as *inferred explicit*. If no, or multiple, clearly stated definitions were found, we searched the article for passages that provided evidence about the definition of productive struggle that might have been used in the study. We then looked across these collected passages for commonalities and selected the passage that best captured those commonalities.

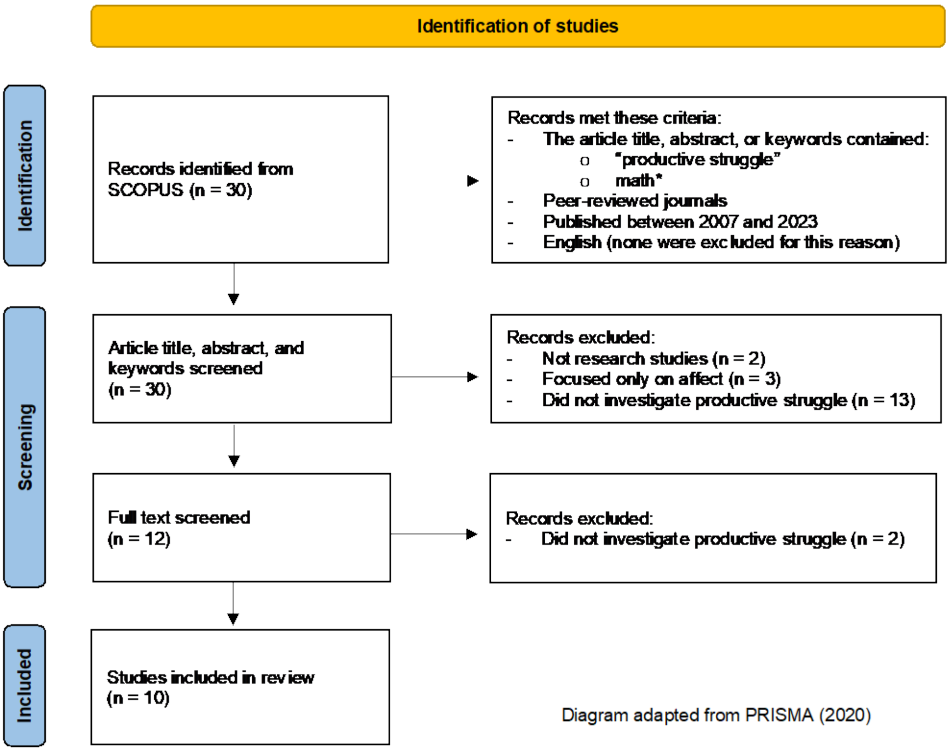



Figure 1: Process of identifying studies included in the systematic review.

This selected passage served as a proxy for the study's definition of productive struggle and was identified as *inferred*. See Figure 2 for a summary of the codes and their definitions.

| Code Name | Code Definition |
|---|---|
| Explicit | The authors stated clearly what definition of productive struggle they used in their study. |
| Inferred Explicit | The authors cited an existing definition that appeared to be the definition of productive struggle they used in their study, but did not state that it was. |
| Inferred | The authors (1) cited multiple existing definitions without indicating which one they used; or (2) only gave descriptions that provided insight into how they might be defining productive struggle in their study. |

Figure 2: Code names and definitions for analyzing productive struggle definitions

To increase the trustworthiness of our review, a mathematics education graduate student researcher familiar with productive struggle double-coded the data for eight of the articles (80%) with the first author. After individually coding, this researcher and the first author compared their codes and through discussion identified the passages that best captured the definition of productive struggle being used in the article. The two authors reviewed the results, revisiting the articles when any questions arose, and agreed on a (proxy) definition for each study, which was reviewed and confirmed by the third researcher.

For (b), the part of our analysis where we deconstructed the structural elements of each definition of productive struggle, the unit of analysis shifted from the study to the identified (proxy) definition. The first author analyzed the (proxy) definitions and identified structural elements that appeared within them. The second author made minor adjustments to this deconstruction and the third researcher verified the accuracy of the results. Finally, for (c) the first author synthesized the structural elements of the 10 productive struggle (proxy) definitions to develop themes across them. The second author and third researcher checked the validity of the synthesis arguments and the strength of the supporting evidence the first author provided.

## RESULTS & DISCUSSION

Figure 3 shows the (proxy) definitions of productive struggle identified in the 10 studies included in our systematic review. Although we were able to identify a (proxy) definition for each study, only 30% of the articles in our study included explicit definitions of productive struggle (W15, W21, W23).

| Authors (ID) | Year | (Proxy) Definition of Productive Struggle | Page | Code |
|---|---|---|---|---|
| Warshauer (W15) | 2015 | "By students' productive struggles, I refer to a student's 'effort to make sense of mathematics, to figure something out that is not immediately apparent' (Hiebert [&] Grouws, 2007, p. [387])" | 376 | Explicit |
| Warshauer et al. (W21) | 2021 | "By productive struggle, we mean what occurs when 'students expend effort in order to make sense of mathematics, to figure out something that is not immediately apparent' (Hiebert [&] Grouws, 2007, p. 387)" | 89-90 | Explicit |
| Warshauer et al. (W23) | 2023 | "By productive struggle, we refer to Hiebert and Grouws' conceptualization (2007) that 'students expend effort in order to make sense of mathematics, to figure out something that is not immediately apparent' (p. 387)" | 3 | Explicit |
| DiNapoli and Miller (D22) | 2022 | "…, Hiebert and Grouws (2007) defined productive struggle as 'effort to make sense of mathematics, to figure something out that is not immediately apparent' (p. [387])" | 2 | Inferred Explicit |
| Zeybek (Z16) | 2016 | "Hiebert and Grouws (2007) [defined] struggle as an intellectual effort students expend to make sense of mathematical concepts that are challenging but fall within the students' reasonable capabilities" | 396 | Inferred Explicit |
| Aljarrah and Towers (A22) | 2022 | "...iterative cycles of 'expressing, testing, and revising mathematical interpretation–and of sorting out, integrating, modifying, revising or refining clusters of mathematical concepts from various topics within and beyond mathematics' ([Lesh & Zawojewski, 2007,] p. 782)" | 857 | Inferred |
| English et al. (E23) | 2023 | "We identified a struggle as resolving productively if it leads to the student's (a) reflection on the limits of his or her previously established knowledge and ability (English, 2013) and (b) perseverance with the activity towards understanding while remaining cognitively engaged in a challenging task (Warshauer, 2015)" | 5 | Inferred |
| Granberg (G16) | 2016 | "A successful, productive struggle would result in the restructuring of mental connections in more powerful, useful ways through which the problem at hand would make sense and new information, ideas and facts would become assimilated (Hiebert & Grouws, 2007)" | 34 | Inferred |
| Rahman (R23) | 2023 | "Productive struggle is when students persevere through challenging tasks leading to mathematical understandings (VanLehn et al., 2019[*])" | 113 | Inferred |
| VanLehn et al. (V21) | 2021 | "...productive struggle [is when students] work hard…to solve challenging, open-ended problems that afford many mathematical insights and discussions" | 994 | Inferred |

*Note that the VanLehn et al.'s (2019) article cited here is an online first version of VanLehn et al.'s (2021) article.

Figure 3: (Proxy) definitions of productive struggle in math education research

The lack of an explicit definition leaves open the possibility of misinterpretation, makes it difficult for researchers to build on each other's work, and inhibits progress in the field. Note that 80% of the articles in our study cited Hiebert and Grouws' (2007) *struggle* definition (all except A22, V21) and all three studies with an explicit definition (W15, W21, W23) used Hiebert and Grouws' (2007) definition for *struggle* as their definition for *productive struggle*. In addition, both of the inferred explicit proxy definitions referred to Hiebert and Grouws' (2007) *struggle* definition. This use blurs the line between *struggle* and *productive struggle*. While all of these articles include descriptions of productive vs. unproductive struggles, it is worth thinking about whether that difference should be clearly articulated by explicitly defining what is struggle, what is productive struggle, and what makes a struggle productive. Doing so would help minimize the possibility of misinterpretation and make it easier for researchers investigating productive struggle in mathematics learning to build on each other's work.

Our deconstruction of the structures within each (proxy) definition of productive struggle revealed four distinct structural elements: (a) the **subject**, (b) the **action**, (c) the **object** the subject needs to do that action, and (d) the **aim** of that action. Figure 4 shows the structural elements for each definition in our study.

| ID | Subject | Action | Object | Aim |
|---|---|---|---|---|
| W15 | student's | effort | mathematics, something that is not immediately apparent | to make sense of, to figure…out |
| W21 | students | expend effort | mathematics, something that is not immediately apparent | in order to make sense of, to figure out |
| W23 | students | expend effort | mathematics, something that is not immediately apparent | in order to make sense of, to figure out |
| D22 | - | effort | mathematics, something that is not immediately apparent | to make sense of, to figure…out |
| Z16 | students | expend [effort] | mathematical concepts that are challenging but fall within the students' reasonable capabilities | to make sense of |
| A22 | - | expressing, testing, and revising…sorting out, integrating, modifying, revising or refining | mathematical interpretation…clusters of mathematical concepts from various topics within and beyond mathematics | - |
| E23 | student's | reflection, perseverance | on the limits of his or her previously established knowledge and ability, the activity… a challenging task | towards understanding |
| G16 | - | reconstructing | mental connections | [the problem] would make sense and new information, ideas and facts would become assimilated |
| R23 | students | persevere | challenging tasks | leading to mathematical understandings |
| V21 | students | work hard | challenging open-ended problems | to solve…afford many mathematical insights and discussions |

Figure 4: A deconstruction of the structural elements of each (proxy) definition of productive struggle

We share here two observations from our synthesis of the information in Figure 4: differences in the definition foci and notable features of the objects. First, we noticed a difference between the (proxy) definitions that had the action of “effort” and the aim “to make sense of” (W15, W21, W23, D22, Z16) and the proxy definition for G16, which gave a more specific action of “reconstructing” and suggested that this action would lead to sense being made and new information, ideas and facts becoming assimilated. This led us to wonder about the advantages of having a broad definition of productive struggle versus specifying actions that students are expected to do when engaging in productive struggle, as did G16 and A22. We also wondered about what exactly made a struggle productive: the opportunity to better understand a mathematical idea or better understanding that idea. We can see advantages either way, but it seems important to state clearly which approach a study has taken. Second, we noticed two notable features of the objects: challenging and within reach. We interpreted “something that is not immediately apparent” (W15, W21, W23, D22) as synonymous with “challenging” (Z16, E23, R23, V24). Thus, 80% of the articles in our study had challenge as a part of their (proxy) definition of productive struggle. In contrast, only one article specified that the objects (e.g., task, mathematical concept) should “fall within the students’ reasonable capabilities” (Z16). Given that Hiebert and Grouws (2007) described Vygotsky’s (1978) zone of proximal development as “the space within which a student’s struggle is likely to be productive” (p. 388), we wondered whether the object of the struggle “being within reach” is just as important to a definition of productive struggle as the object being challenging.

## CONCLUSION

To enable the field to rethink together how we can investigate what it means for mathematics learners to engage in productive struggle, we shared our preliminary systematic literature review about how *productive struggle* is defined in mathematics education research. Our findings provided insight into the explicitness, coherence, and variation of current definitions. They also highlighted the importance of all researchers explicitly identifying the key terms they use in their work. Doing so would support moving the field forward as it would increase the ability of researchers to accurately build on each other’s work and clarify their contributions to the field.

We limited our research to studies investigating productive struggle in mathematics learning published in peer-reviewed journal articles included in the Scopus database. Doing so potentially excluded valuable insights. For example, Kapur’s definition of *productive failure* as “the design of conditions for learners to persist in generating and exploring representations and solution methods (RSMs) for solving complex, novel problems” (Kapur, 2021) seems directly related to productive struggle. Similarly, Sengupta-Irving and Agarwal’s (2017) study was excluded from our study because it focused on perseverance rather than productive struggle. However, they explicitly defined productive struggle using Hiebert and Grouws’ (2007) definition of struggle as part of their explicit definition of *perseverance in problem solving as collective enterprise.* This idea of "collective enterprise" seems worth considering as the (proxy) definitions in our study did not directly address the potentially collective nature of productive struggle (see, for example, Kamlue & Van Zoest, 2022). These two examples illustrate why it might be useful to expand our literature review. An expanded systematic literature review could include (a) aspects of research on productive struggle that we excluded, (b) other types of publications (e.g., dissertations, book chapters, and conference proceedings), and (c) other fields that are related to mathematics education. The first and second expansions would provide a more comprehensive look at how productive struggle is defined in mathematics education research, and the third would provide insight into how those in other fields are using the term. Future research could use our methodology and findings to sharpen the construct of productive struggle for the field of mathematics education.

**Acknowledgement**

We thank Yaronn James Arciaga for increasing the trustworthiness of our study.

**References**

References marked with an asterisk indicate studies included in the systematic review.

*Aljarrah, A., & Towers, J. (2022). The Emergence of Collective Mathematical Creativity Through Students' Productive Struggle. *Canadian Journal of Science, Mathematics and Technology Education*, *22*(4), 856-872.

*DiNapoli, J., & Miller, E. K. (2022). Recognizing, supporting, and improving student perseverance in mathematical problem-solving: The role of conceptual thinking scaffolds. *The Journal of mathematical behavior*, *66*, 1-20.

*English, A. R., Tyson, K., Hintz, A., Murdoch, D., & Anderson, J. (2023). Pedagogical listening: Understanding how teachers listen to student struggle during mathematical sense-making discussions. *Teachers and Teaching*. Advance online publication. https://doi.org/10.1080/13540602.2023.2263738

*Granberg, C. (2016). Discovering and addressing errors during mathematics problem-solving—A productive struggle? *The Journal of Mathematical Behavior*, *42*, 33-48.

Hiebert, J., & Grouws, F. A. (2007). The effects of classroom mathematics teaching on students learning. In J. Frank, & K. Lester (Eds.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 371–404). Information Age Publishing.

Kaiser, G., & Schukajlow, S. (2023). Literature reviews in mathematics education and their significance to the field. *ZDM–Mathematics Education*. Advance online publication. https://doi.org/10.1007/s11858-023-01541-z

Kapur, M. (2010). Productive failure in mathematical problem solving. *Instructional Science, 38*(6), 523-550.

Kapur, M. (2021). *Productive failure*. Manu Kapur. https://www.manukapur.com/productive-failure/

Kamlue, N., & Van Zoest, L. R. (2022). A preliminary look at prospective mathematics teachers' productive struggle. In A. E. Lischka, E. B. Dyer, R. S. Jones, J. N. Lovett, J. Strayer, & S. Drown (Eds.), *Proceedings of the 44th PME-NA* (pp. 1206-1210). Middle Tennessee State University.

National Council of Teachers of Mathematics (NCTM). (2014). *Principles to action: Ensuring mathematical success for all*. Author.

Nieminen, J. H., Bagger, A., Padilla, A., & Tan, P. (2023). Student positioning in mathematics assessment research: A critical review. *Journal for Research in Mathematics Education*, *54*(5), 317-341.

Phan, T. T., Do, T. T., Trinh, T. H., Tran, T., Duong, H. T., Trinh, T. P. T., ... & Nguyen, T. T. (2022). A bibliometric review on realistic mathematics education in scopus database between 1972-2019. *European Journal of Educational Research*, *11*(2), 1133-1149.

PRISMA: Transparent Reporting of Systematic Reviews and Meta-Analyses (PRISMA). (2020). *Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)*. Home. http://www.prisma-statement.org/

*Rahman, Z. G. (2023). Pre-service mathematics teachers' experience with productive struggle. *The Educational Forum 87*(2), 112-130.

Scopus blog. (February 15, 2023). *Scopus now includes 90 million + content records!* Scopus. https://blog.scopus.com/posts/scopus-now-includes-90-million-content-records

Sengupta-Irving, T., & Agarwal, P. (2017). Conceptualizing perseverance in problem solving as collective enterprise. *Mathematical Thinking and Learning, 19*(2), 115-138.

Warshauer, H. K. (2011). *The role of productive struggle in teaching and learning middle school mathematics* [Doctoral dissertation, University of Texas at Austin]. Texas ScholarWorks. https://repositories.lib.utexas.edu/server/api/core/bitstreams/ca5906d0-c7d 4-4e75-9f3c-6e1fac54d94e/content

*Warshauer, H. K. (2015). Productive struggle in middle school mathematics classrooms. *Journal of Mathematics Teacher Education, 18,* 375-400.

*Warshauer, H. K., Herrera, C., Smith, S., & Starkey, C. (2023). Examining preservice teachers' noticing of equity-based teaching practices to empower students engaging in productive struggle. *The Journal of Mathematical Behavior*, *70*. Advance online publication. https://doi.org/10.1016/j.jmathb.2023.101045

*Warshauer, H. K., Starkey, C., Herrera, C. A., & Smith, S. (2021). Developing prospective teachers' noticing and notions of productive struggle with video analysis in a mathematics content course. Journal of Mathematics Teacher Education, 24, 89-121.

*Zeybek, Z. (2016). Productive struggle in a geometry class. *International Journal of Research in Education and Science*, *2*(2), 396-415.

*VanLehn, K., Burkhardt, H., Cheema, S., Kang, S., Pead, D., Schoenfeld, A., & Wetzel, J. (2021). Can an orchestration system increase collaborative, productive struggle in teaching-by-eliciting classrooms? *Interactive Learning Environments*, *29*(6), 987-1005.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Harvard University Press.

# NAVIGATING FLIPPED LEARNING: INSIGHTS FROM A GRADUATE-LEVEL ALGEBRAIC GEOMETRY COURSE

Sang Hyun Kim, Tanya Evans, and Ofer Marmur

University of Auckland, New Zealand

*This study explores the integration of flipped learning into a graduate-level algebraic geometry course, addressing gaps in understanding its implementation at this educational level. Through an exploratory case study, students' experiences were examined, and thematic analysis revealed that students had nuanced perceptions of this integration with four major themes arising: Preparation and Workload, Content Interaction, Social Interaction, and Resources. While students appreciated collaborative aspects and the emphasis on problem-solving, challenges emerged, including an increased workload and a strong preference for explicit forms of instruction. This research underscores the need for further exploration to refine flipped learning practices and gain a comprehensive understanding of its implications on student experiences in graduate mathematics education.*

## INTRODUCTION

Since its emergence, flipped learning has garnered much attention from researchers and practitioners, acknowledged for its ability to enhance inter-student interactions through dynamic and collaborative practices (Bergmann & Sams, 2012). As a mode of instruction, it challenges the traditional classroom norms and supports a more dynamic environment where students take on more responsibility in their learning. In this setting, students are expected to engage with mathematical content before meeting with the instructor and their peers. However, in practice, successful implementations of an instructional mode can be complicated and met with many challenges (Lo et al., 2017). It relies on the responsibilities of teachers and students alike, and more work is needed to understand its implementation at the graduate level.

This study reports on an attempt to integrate aspects of flipped learning into a graduate-level mathematics course. By infusing elements of flipped learning into a traditionally lecture-based course, it was assumed that the limitations of both modes of instruction could be addressed and mitigated. This exploratory case study sought to explore how flipped learning can be realised in a largely untouched context at the highest level of tertiary maths education by considering student perspectives. To this end, we pose the following research question: How do students perceive aspects of flipped learning in a graduate mathematics course?

## FLIPPED LEARNING

Flipped learning is positioned in a broader model of learning: blended learning. Despite numerous definitions posed across the literature over time (Bishop & Verleger, 2013), we adopt a common definition used by the Flipped Learning Network (FLN) (2014),



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 129-136). PME.

which describes it to be 'a pedagogical approach in which direct instruction moves from the group learning space to the individual learning space.' Their framework outlines the four main pillars that must be incorporated into practice before it is considered flipped learning: Flexible Environment – the importance of being able to provide varied ways of engaging with the content; Learning Culture – the active role students have and evaluate their learning; Intentional Content – careful consideration of the material to be covered independently and in class; and Professional Educator – the essential role of the instructor in the classroom to provide feedback, monitor student progress and the practices in the lesson.

The reported effects of flipped learning on learning achievement in mathematics generally demonstrate a positive trend and highlight various advantageous outcomes, including improved opportunities for feedback, interactions, and applying concepts in class (Lo et al., 2017). Similar findings were echoed by Cevikbas and Kaiser (2022), who found that the reported opportunities offered by a flipped classroom approach were often related to conceptual gains, engagement, and collaboration. On the other hand, the challenges of a flipped learning approach reflect what happens within the lessons themselves and outside of them. Cevikbas and Kaiser (2022) report on four main groups of challenges: pedagogical (e.g., lack of preparation by students), technical (e.g., poor internet connectivity), cognitive (e.g., difficulty remembering lecture video content), and affective issues (e.g., lack of motivation). Students' unfamiliarity with the approach and the high investment required by instructors are two further oft-reported challenges (Lo et al., 2017). Flipped learning in mathematics has been researched extensively; however, much insight stems from research at the undergraduate level (Lo et al., 2017). Our exploratory study aims to fill a critical gap in the graduate mathematics literature by exploring the integration of flipped learning into a graduate-level course and providing insights into student experiences and perceptions.

## METHOD

### Research Design

An exploratory case study research design was employed to explore the implementation of flipped learning into our course of interest. This was suitable as this course offered an opportunity to investigate a revelatory case, a phenomenon in a context that had not been explored before (Yin, 2018, p. 50).

### Context

The course within this study was a graduate-level algebraic geometry course at a large research university in New Zealand, with a particular focus on curves. Most iterations of this course see 10-15 students enrolled. The instructor assigned a set of YouTube videos created by a Fields medallist, which followed the structure of a traditional lecture and were around 20 minutes each. The course instructor kept some of the class time for revisiting the concepts and ideas from the videos and usually would spend about 20 minutes (out of 50) providing a more traditional lecture on the whiteboard addressing any complex or critical ideas. The course had various aspects of flipped learning implemented into it: the offering of multiple ways to engage with the content through many resources and activities (Flexible Environment); the greater time students had to engage with problem sheets and discuss the content (Learning Culture); the intentional assignment of resources and the curation of complementary problem sheets (Intentional Content); and the availability of the instructor during the lessons, both to provide feedback and monitor ongoing student progress (Professional Educator).

**Participants and Data Collection**

About half of the students were graduate students, while the other half were undergraduate students who had obtained special permission to enrol in the course. Of the fourteen students in the course, twelve participated in this study. All the students involved were studying mathematics as part of their academic programmes, while many students were also immersed in related fields such as physics, statistics, and finance. All students were sent surveys to complete via Qualtrics on the penultimate teaching week of the second semester of 2023. The survey contained questions about student background, affective factors, course engagement, and various open-ended questions exploring student perspectives of the course. The surveys were conducted to gauge the student perspectives on their experiences on several aspects of the course.

The analysis of this study involves student responses to four open-ended questions: (1) ‘What are the aspects you liked about this course or what aspects did you dislike?’, (2) ‘Regarding the flipped learning aspect of the course, how did you think this affected your experience for the course?’, (3) ‘What would you like to see improved in a course like this?’, (4) ‘What is your preferred learning model you prefer and why? (e.g., flipped lectures, traditional lectures, online lectures, a mix, etc.)’. The open responses allowed participants to elaborate on more nuanced aspects of their experiences than can be seen by closed-response questions alone.

**Analysis**

Thematic analysis was used to identify relevant themes from the open response data. The thematic analysis approach used in this study was inductive, where themes and codes were driven from the data rather than selected a priori. A thematic analysis provides a helpful way to view qualitative data, uncover new perspectives, and identify the similarities and differences between participant responses.

We engaged in a complete coding process where the open-response survey data from all twelve participants were read, and all potential features of the data were coded before engaging in further iterative processes. While we have tried to avoid actively construing meaning in a way that is unsupported by the responses, there is a degree of interpretation required in selecting codes and themes. Our active position as researchers within this process cannot be avoided, and we recognise the inevitable ever-present bias in any form of research. Any codes irrelevant to addressing our research question were omitted from the analysis.

## FINDINGS

The comments directly related to the flipped aspects of the course were nuanced, with different dimensions and preferences being voiced by various students. Participants frequently provided balanced views in these comments; however, some clearly articulated dichotomous comments—either positive or negative—regarding the experience. Many comments were quite assertive, with a few stating that flipped learning does not work for a demanding course. As one student said, ‘It was a difficult course unfortunately and thus required better instruction than could be provided from flipped lecture’ (student 4). Others noted that the experience was nice, unexpected, and comparable to traditional modes of instruction in difficulty. From the thematic analysis, four major themes regarding how students perceive flipped learning aspects in the course were identified from the data: *Preparation and Workload*, *Content Interaction*, *Social Interaction*, and *Resources*. In the following section, each theme is explained, and extracts illustrating themes are provided.

### Preparation and Workload

The theme of *Preparation and Workload* captured codes related to one of the major differences between a flipped classroom and a more traditional instructional approach, which is the work required by students to prepare for classes. Students are required to engage with content before a lesson, unlike a traditional lecture where content engagement typically happens for the first time during the lecture.

Students attributed the higher-than-usual workload to the preparation needed to be done for the lessons, and sometimes this increase was perceived to be significant: ‘The current set-up with flipped lectures did dramatically increase the workload of the course by having recordings to watch outside of the lecture times’ (student 8). Additionally, this preparation was seen by some students as being essential to participating or even simply attending the lectures. For some, a lack of proper preparation meant it would be ‘very easy to fall behind’ (student 4). One participant noted that they ‘Missed out much of the opportunity to practice solving problems when fallen behind on lectures’ (student 7). For others, it was the case that preparation, or the lack thereof, impacted their attendance. One student expressed their views as such: ‘I was more motivated to keep up to date with the lecture content. When I haven't kept up to the content, I was less likely to show up to the lectures’ (student 7). This student reported that the flipped learning aspect had encouraged them to stay up to date with the current workload but that when they slipped behind, they felt less inclined to attend the lectures. Falling behind did not seem uncommon among the participants, which is not ideal considering that the flipped aspects were opportunities for students to apply the knowledge and skills they have encountered in different contexts. Furthermore, with many concepts being so highly connected, it is unlikely that this will have no bearing on a student’s ability to engage with later content. Unfortunately, this appears to have been the case for one student who reported this to be a problem for them across the semester: ‘I was not able to consistently watch the recordings due to the increased work load of the course, and as a result I ended up falling behind in the content for most of the course' (student 8).

**Content Interaction**

*Content Interaction* concerns comments coded for relevance to how students engage with the learning material. This may be through the tasks during class, such as, most often mentioned, solving exercises in class, but comments under this theme also reflect the difficulties students had in keeping up with more demanding content and a desire to revise fundamental knowledge.

The opportunity to engage with the content by completing exercises during class time, something strongly associated with flipped classrooms, seemed to resonate with individuals who liked being able to: '(…) ask questions and actually doing exercises in class' (student 1). Other participants shared appreciation for this aspect, but one student expressed that some more variety in content engagement would have been helpful: "I liked the emphasis on solving exercises, but I think I may have benefitted if there was more proving theorems and discussion fundamental notions in class" (student 9). This comment may have referred to a desire for more opportunities to observe the lecturer demonstrating these skills in the lessons, more than was offered in the course. For instance, more explicit demonstrations on the whiteboard and during lectures (e.g., students 4 & 9). This desire may have resulted from the high complexity and novelty of the mathematics within the course where students may require more guidance and explicit instruction. Likewise, it may be that the inherent complexity observed in a graduate-level course, along with the fast pace, 'Made it harder to quickly build understanding of the content as opposed to being able to actively engage with the lecturer as they explain the content' (student 2). This comment could be comparing the decrease in explicit instruction and the instructor's demonstration, which is more typically seen in traditional lectures.

**Social Interaction**

The final theme was that of *Social Interaction*. This encompassed codes about opportunities for interaction with peers and the instructor that were afforded by integrating flipped learning approaches in the lessons. This theme may be unsurprising, as the emphasis on collaborative practices was a noteworthy change for students compared to their other mostly traditional style courses. There is a notable variation in the codes within this theme, highlighting students' preferences and a need to consider these in any setting. A few students commented favourably about the engagement with peers and instructors in the course: 'I liked (…) the engagement with peers and with the lecturer' (student 5); however, most students did not explicitly discuss this in their responses. Regarding the learning atmosphere, one student described flipped learning as an environment conducive to asking questions; 'There are [sic] adequate time for ask [sic] questions' (student 1). On the other hand, the flipped nature may have incited more questions, which students felt could not be addressed or attended to. The following excerpt illustrates this:

> ‘Flipped lectures are ok, because they provide dynamic feedback and seem to foster peer collaboration, however I still have a tendency to get lost as the pace is still fast, and I am not able to clarify all my confusions (this would entail me asking a question every 10 seconds so would be unreasonable).’ (student 5)

When comparing the current course with other more traditional mathematics courses in the past, a similar view was expressed about the preference for traditional lectures in addressing clarifications over the current flipped format. Perhaps, similar to the previous student excerpt, the pace of the course, instructor availability, or another factor may have made it more challenging to engage in such forms of interactions:

> ‘I think the majority of the classes should still be taught in the traditional lecture format though as it is much easier to clarify something that you don't understand in this setting, as opposed to going through flipped lectures, making a list of what you don't understand and then having to e-mail the lecturer/attend office hours etc.’ (student 11)

On a similar note, student 4 states that in more traditional settings, ‘it just feels like you can ask questions as you're learning and not leave out any holes right at the beginning.’

**Resources**

The final theme encompasses codes directly linked to the resources students were provided with and their experience engaging with them outside of class. This theme is important to consider as one notable aspect of flipped learning is providing students with an adequate opportunity to appropriately engage with content in a way that best assumes the role of traditional lectures. This would include ensuring it is easy to use and navigate between resources.

One student notes the abundance of ‘all available resources (videos, flipped classes, notes, textbooks)’ (student 7) as a benefit. However, it is not hard to imagine that providing many resources can just as easily be experienced as a burden, especially if a student believes they lack the necessary background knowledge, making it hard to know how to use various resources. To mitigate this potential source of difficulty, student 6 suggested that it ‘Would've been better if we followed only one [set of videos] videos or [the textbook]’ (student 6). The diversity of resources resulted in inconsistencies in the definitions of key concepts: ‘The lectures often taught things different to [the textbook] which meant people had different definitions for things and it was difficult to reconcile these’. Using external resources is a defining feature of any flipped learning experience (Flipped Learning Network (FLN), 2014) and one that will depend on when students work with more complex mathematical concepts.

## DISCUSSION

This study aimed to explore the flipped learning facet, an unfamiliar approach to learning mathematics for many students, introduced to a graduate-level mathematics course and how students perceived it. The thematic analysis identified four key themes that students reported, which closely align with many of the key characteristics of flipped classrooms. Students' experiences within this graduate-level course are parallel to similar studies, suggesting that there may be more commonalities between implementing flipped learning than we think. For instance, the emphasis on problem-solving in flipped learning and opportunities for collaboration with peers are commonly lauded (Lo et al., 2017). Additionally, a notable increase in student workload (in *Preparation and Workload)* is a widely reported consequence of realising flipped classrooms in the literature (Cevikbas & Kaiser, 2022), one which can result in students falling behind and being unable to participate.

The participants in this study showed a preference for traditional instructional modes, with some students finding it easier to seek clarifications and ask questions in lectures than in a flipped classroom. Moreover, the difficulty of the course and the demanding nature of the content may have also shaped this perception. Previous research suggests that an inclination for lectures may be due to the greater structure of lessons or the higher levels of reported concentration during them (Feudel & Fehlinger, 2023). Additionally, Novak et al. (2017) state that while flipped learning may be appropriate for the development of practical skills, the introduction of concepts may be better suited to lectures where a more knowledgeable figure can better support mathematics learning through explicit instruction (for review, see Evans & Dietrich, 2022). This is similar to mathematicians who often lean towards lectures as a preferred mode of learning from their colleagues as a means to support engagement with new mathematical areas (Weber & Fukawa-Connelly, 2023). Similarly, high-achieving students may share this sentiment for lectures as they encounter novel mathematics.

This study contributes to the graduate mathematics literature by providing one of the first studies to report on implementing flipped learning at this level. By doing so, we have shown that certain aspects can be well-received by students while others require further consideration. Implementing the main pillars of flipped learning may be more intricate at the graduate level, and this study paves the way for future research to explore such avenues. The limitations of a small-scale exploratory case study are that it is difficult to generalise our findings. The insights from this study contribute to the broader conversation on instructional approaches in graduate mathematics education and the student experiences shaped through them.

## References


Bergmann, J., & Sams, A. (2012). Flip your classroom: Reach every student in every class every day. International society for technology in education.

Bishop, J., & Verleger, M. A. (2013). The flipped classroom: A survey of the research. 2013 ASEE Annual Conference & Exposition,

Cevikbas, M., & Kaiser, G. (2022). Can flipped classroom pedagogy offer promising perspectives for mathematics education on pandemic-related issues? A systematic literature review. ZDM – Mathematics Education. https://doi.org/10.1007/s11858-022-01388-w

Evans, T., & Dietrich, H. (2022). Inquiry-based mathematics education: a call for reform in tertiary education seems unjustified. STEM Education, 2IS - 3SP - 221, 244. https://doi.org/10.3934/steme.2022014

Feudel, F., & Fehlinger, L. (2023). Using a lecture-oriented flipped classroom in a proof-oriented advanced mathematics course. International Journal of Mathematical Education in Science and Technology, 54(1), 46-73. https://doi.org/10.1080/0020739X.2021.1949057

Flipped Learning Network (FLN). (2014). The Four Pillars of F-L-I-P™.

Lo, C. K., Hew, K. F., & Chen, G. (2017). Toward a set of design principles for mathematics flipped classrooms: A synthesis of research in mathematics education. Educational Research Review, 22, 50-73. https://doi.org/https://doi.org/10.1016/j.edurev.2017.08.002

Novak, J., Kensington-Miller, B., & Evans, T. (2017). Flip or flop? Students' perspectives of a flipped lecture in mathematics. International Journal of Mathematical Education in Science and Technology, 48(5), 647-658. https://doi.org/10.1080/0020739X.2016.1267810

Weber, K., & Fukawa-Connelly, T. (2023). What mathematicians learn from attending other mathematicians' lectures. Educational Studies in Mathematics, 112(1), 123-139. https://doi.org/10.1007/s10649-022-10177-x

Yin, R. K. (2018). Case study research and applications : design and methods (Sixth edition. ed.). SAGE.

# STUDENTS' VIEWS OF E-ASSESSMENT FEEDBACK IN UNDERGRADUATE MATHEMATICS

George Kinnear and Paola Iannone

The University of Edinburgh

*This paper reports on undergraduate mathematics students' views on the feedback delivered through an e-assessment system, based on thematic analysis of interviews with 20 students. The results highlight students' views on the content of feedback – with many students expressing a preference for detailed, specific feedback, and mixed opinions about whether e-assessment delivered this. Students also reported strategic approaches to using the feedback. The findings resonate with existing frameworks on students' interactions with feedback, and provide a basis for further work to explore students' views toward e-assessment feedback in other contexts.*

## INTRODUCTION

Undergraduate mathematics education has increasingly embraced e-assessment as a tool for both formative and summative purposes (Kinnear, Jones, et al., 2022). Much previous research has investigated students' views about assessment, since these views influence how students approach their studies (Van der Kleij & Lipnevich, 2020). For instance, Iannone and Simpson (2015) found that mathematics students expressed a preference for more traditional forms of assessment, including closed-book exams, in contrast to findings in other disciplines. However, students' views about e-assessment have so far received relatively little attention.

Researchers studying the use of e-assessment within undergraduate mathematics have begun to explore students' views, and how these interact with students' activity. In Norway, Rønning (2017) surveyed engineering mathematics students and found that they appreciated getting feedback immediately on their weekly e-assessments – although they were not satisfied with the quality of the feedback, which gave "no indication of the reason for an error" (p. 96). This sometimes led to students adopting an approach of "hunting for the answer" (p. 101). In the US, Dorko (2020) combined observations of students working on e-assessment tasks with follow-up interviews, to develop a model of students' engagement. This study highlighted the cyclic nature of students' work on e-assessment tasks, particularly where multiple attempts were allowed, since students used initial attempts as "formative assessment to inform their work on the rest of a problem" (p. 463).

These studies highlight the importance of the way that students view feedback from e-assessment. While there is a large body of work on student perceptions of feedback (for a review, see Van der Kleij & Lipnevich, 2020), very little of this has focused on e-assessment feedback, and less still on undergraduate mathematics.



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 137-144). PME.

Our research question for this study was: what are students' views toward e-assessment feedback in an undergraduate mathematics course?

## THEORETICAL BACKGROUND

Van der Kleij and Lipnevich (2020) reviewed 164 studies on students' perceptions of feedback. They summarised the findings about factors that influenced students' views under three broad headings: characteristics of the feedback, characteristics of the student, and the student response to the feedback. For our study, we did not seek to investigate whether characteristics of the students (e.g., gender, age, or other demographic variables) moderated their perceptions; the summary of previous research indicated studies of those characteristics tended to have findings that were inconclusive, or inconsistent across studies. The other two factors are relevant for our study, and were elaborated on by Lipnevich and Smith (2022) in their proposed model for the entire student-feedback interaction.

Regarding characteristics of the feedback, Lipnevich and Smith (2022) highlighted various aspects, including the timeliness, accuracy, and level of detail. All of these are pertinent to our study of views about e-assessment feedback, where these characteristics may be expected to differ from other forms of assessment (for instance, e-assessment feedback can be delivered more quickly than written feedback).

The student response to feedback is a crucial part of the model proposed by Lipnevich and Smith (2022). Their model emphasises students' generation of "self-feedback" in response to feedback messages, which can be described using three questions: "Do I understand the feedback? How do I feel about the feedback? What am I going to do with the feedback?" (p. 3). These three questions capture respectively the cognitive, affective and behavioural responses to feedback, which "interact with one another and result in self-feedback that defines action – or in a decision not to do anything about the feedback." (Lipnevich & Smith, 2022, p. 3).

Through their review, Van der Kleij and Lipnevich (2020) highlighted that interviews were a commonly-used method for research on this topic (64 out of 164 studies), although a majority of those had "serious methodological violations" (p. 358), which we sought to avoid in our study.

## METHOD

We carried out semi-structured interviews with 20 students studying the same first-year pure mathematics course at a research-intensive UK university. The course is one third of a normal load for the semester, notionally requiring 200 hours of study across the 11-week semester. The course includes weekly written assignments that are marked by tutors (together counting for 25% of the course grade), and weekly e-assessment quizzes delivered through the STACK e-assessment system (for another 25% of the course grade). The e-assessment quizzes feature novel task types, including proof comprehension tasks (as described by Bickerton & Sangwin, 2021).

The interview questions aimed to elicit students' views about various aspects of e-assessment; one planned question that was particularly relevant to the analysis here was: "what do you think of the feedback that you get from STACK?". Participants were also prompted to make comparisons with the traditional written assessments in the course. All students on the course were invited to participate, and 20 students responded. We interviewed the students in March 2022, with 11 students in person (assigned pseudonyms beginning with P) and 9 online (assigned pseudonyms beginning with S). All interviews were recorded and later transcribed.

For the analysis, we used a thematic analysis approach (Braun & Clarke, 2006). We used descriptive coding in the first round of analysis, and developed themes and sub-themes from these. We revisited all transcripts to check that the themes were representative and identify relevant passages that had been missed in the first round.

## RESULTS

The overall thematic map is shown in Table 1. For reasons of space, the analysis we report here is focused on the "Response to feedback" theme only.

| Theme | Sub-themes |
|---|---|
| Approach to assessment | • Use of time<br>• Work with others<br>• Activity (reading, writing, watching lectures) |
| Response to feedback | • Content of the feedback<br>• Using the feedback<br>• Focus on marks<br>• Timing of the feedback |
| Appreciation of the assessment | • Emotions<br>• Learning and study approaches<br>• Developing mathematical skills |

Table 1: Themes and sub-themes developed from the interviews.

### Content of the feedback

Students consistently expressed a preference for feedback that is specific and detailed, and accordingly tended to favour the feedback that they got on written work over the e-assessment feedback. However, several students noted that detailed feedback on written work was not provided consistently. For instance, S1 said that "I don't think I've ever got particularly detailed feedback for the hand-ins", and that they found the e-assessment feedback helpful because "it'll usually give a solution". S8 also appreciated the worked solutions provided in the e-assessment feedback, since "you can see how it's meant to be done", in contrast with feedback on written work which is "just sorta saying where you went wrong and … not much like, well, feedback on what you need to change”. Many of the students also commented favourably on the depth of explanation in the e-assessment feedback. For instance, P11 described the feedback as “really helpful” and highlighted how in solutions to multiple-choice questions “it talks about each case … tells you, oh yeah, this one’s right, this one’s wrong, here's why it’s both of them … it’s just really helpful with the detail”.

Many students expressed dissatisfaction with e-assessment feedback that was limited to a score and “just a kind of generic solution - here’s how to solve this question” (S1). Indeed, P6 characterised this as “not really feedback”. Many students compared this feedback to feedback they received on written work, which is “actually marked by people, so they can point out specific things” (P11). S7 explained that tutors can “pinpoint exactly where we went wrong and they can, like, suggest a better solution to it”, whereas the e-assessment feedback “can’t tell me how I went wrong”. In fact, the e-assessment system does have the facility to include specific comments based on the student answer, provided the teacher has programmed this in advance (for an example, see Alarfaj & Sangwin, 2022). P11 recalled receiving some feedback like this: “sometimes it will say, oh yeah, this wouldn’t work because of such and such … I don’t know if it does that all the time, but sometimes it will, and it’s very helpful”. Similarly, P10 said that e-assessment feedback will “most of the time … tell you where you’ve gone wrong”, although “there have definitely been sometimes where even after reading it, I’ve maybe been like, so where did I go wrong then?”.

Some students identified other drawbacks of the e-assessment feedback being largely limited to a score and a model answer. P7 noted that sometimes the model answer was “repeating something I guess I could read in the textbook” which was not helpful if that explanation “didn’t click”. S6 found that the model answers could help them see the correct answer but “it doesn’t necessarily help me from like an exam perspective … because it doesn’t really tell me how exactly I should be writing down the solution”; they also described how their tutorial classes would often include discussion of the relative merits of different solution methods, and “that kind of feedback is missing” in the e-assessments.

**Using the feedback**

Several students emphasised the independence required to process and respond to the e-assessment feedback when it took the form of a model solution (which highlights the connection between this sub-theme and the “content of the feedback” sub-theme). For instance, P7 described how:

> there are some questions where the feedback doesn’t help at all … I’ve no idea why my version isn’t correct if I’ve done it in a slightly different way … [the e-assessment feedback] will just tell me I’m wrong, and I will have to figure out why myself, or what I need to fix myself.

This independent sense-making is more difficult when the model solutions omit steps in the working: S1 noted that “it might take me a while to figure it out if I didn’t really understand how to do the question to begin with, then sometimes I’ll find there’s steps they're skipping that I'm not totally sure about". Nevertheless, it seemed that many students persisted in their efforts: for instance, P8 described how "at first I throw my hands up and go, great, but then I probably just go over the notes again and try to understand, yeah, make sense of it in the context of the notes".

Other students described using the feedback for more general purposes, by extrapolating beyond the individual tasks that the feedback was addressing. P8 said they used the e-assessment feedback to identify topics they needed to revisit, and noted that getting things wrong "helps you then reflect more deeply on stuff you've learned". Similarly, P10 valued the e-assessments providing a means "to be able to actually know where we're at, to make progress, to ask questions, to know what we need to read more into, to just, like, develop our understanding".

Students also described making decisions about when to seek further help, from peers or their teachers, on the basis of the feedback. P10 said that when the e-assessment feedback presented a model solution based on a different method to the one they had used, they approached friends on the course who had used the same method to compare working and find the step where they had gone wrong. P11 said that after reviewing the feedback, "if you feel that you're not doing well, then you can bring it up with people, like, ask tutors about it". However, students were not always proactive in seeking help; for instance, P5 described finding the e-assessment feedback hard to understand, and said that "sometimes I ask the tutor in the tutorial, but other times, I just ignore".

Many students described using the e-assessment feedback in a strategic way. One strategy was specific to e-assessments where multiple attempts were allowed: some students described using an initial attempt as a means to see the worked solutions, which they could then use to guide their work on an actual attempt at the quiz. For instance, P6 described entering a "random answer, just so I can get that yellow box" with the worked solution. Similar behaviour has been identified in previous research on students' use of e-assessment (Kinnear, Wood, et al., 2022), where it was described as "gaming the system".

Another way that students were strategic was through engaging selectively with the feedback, including making choices about whether or not to act on it. For instance, P7 described reviewing the feedback to "look through all the ones I got wrong", but that they "don't normally go back immediately and fix the stuff … my plan is to go back once I've finished my lectures and tutorials and use those quizzes for revision". P9 said that they would look at their score after the deadline for the assessed quiz, but would rarely seek out the feedback, because "it's a hassle to go back, it's more work … with everything else going on". P11 shared an example of a recent e-assessment where they "didn't think the answer … that was given, was actually right", and that they "checked it with someone else who's in second year and they actually agreed with me". When asked what they should do in that situation, they said they "don't really know", so this situation was ultimately unresolved.

**Focus on marks**

Related to the selective engagement with feedback described above, several students were particularly focused on the mark, rather than any other features of the feedback. P7 described checking the feedback when it is released, to "check what I got", and "if it's a much lower mark than I expected, I'll scan down to see what sort of things", while S5 said "if I've got over 80%, it's generally nothing to worry about".

One aspect of this focus on marks that caused particular tension with e-assessment was the students' desire for partial credit. Students highlighted instances where they had made an error but had much of the working correct; for instance, P9 noted that "you could just misinterpret a question, have something that's say double the number that they expect, and you get zero points, whereas on a hand-in, you might've gotten partial marks". Another common issue is where the answer was expected in a particular form: S2 said entering answers in the right format had "been problematic a lot", giving the example of "when integrating and adding a constant … do you have to do a capital C or a lower case C? And people would just like completely get the question wrong cause they did a capital instead of lower case". While several students were frustrated by losing marks over syntax errors, S5 was more relaxed: "I always tend to find that you're not necessarily getting the understanding incorrect … if I've got a question wrong because I've entered a number wrong, I won't worry about it".

**Timing of the feedback**

Many students commented on the timing of feedback from e-assessment, particularly in light of different settings used for two types of quizzes in this course: the formative "lecture quiz" gave immediate feedback while the summative "assessed quiz" only gave feedback after the deadline had passed. Most students expressed a preference for immediate feedback, while the question was still fresh in their mind, and noted that turnaround times on written assignments were typically much longer than for the e-assessments. S8 said that on the lecture quiz, "if I get the question wrong then I am thinking about the question, so I'm much more likely to look at the feedback", while S9 said that for the assessed quiz:

> by the time I get to review my answers, it'll be a couple of days after I did it, and I've sort of lost, like, a track of why I wrote what I wrote, I just know that I got it wrong, so it doesn't, like, help me learn as much.

For this reason, S1 completes STACK quizzes near the deadline, so "you can get the feedback the next day, which is nice". While most students preferred immediate feedback, P7 said they disliked assessed quizzes in another course that gave feedback on each question as it was completed, because it "probably negatively affects my performance on later questions if I know I've already got a few ones beforehand wrong".

## DISCUSSION

We interviewed 20 undergraduate mathematics students about their views on e-assessment, and identified four sub-themes in their response to e-assessment feedback. First, regarding content of the feedback, detailed and personalised feedback was seen as desirable. Second, students described strategic approaches to using the feedback, in light of the effort required to independently process the feedback they received. Third, many students had a focus on marks rather than elaborated feedback, and were dissatisfied with the way that e-assessment often did not give partial credit. Fourth, students expressed a preference for immediate feedback.

Our interviews were designed to explore a range of issues related to e-assessment, not just the views about feedback that we have reported here. More focused interviews may have elicited more detail, and could have explicitly targeted issues that are highlighted in existing frameworks on feedback. On the other hand, our relatively open, semi-structured approach allowed for participants to raise the issues that mattered most to them. While our thematic analysis was inductive, the sub-themes that we identified have nevertheless aligned with aspects of existing frameworks; for instance, our "use of feedback" sub-theme is in line with the emphasis of Lipnevich and Smith (2022) on the way that students engage with feedback.

A limitation of our study is that it took place in one particular context, which will inevitably have shaped students' views. For instance, at this institution, students are used to completing weekly e-assessment tasks that make up a small part of the course grade. In the interviews, students would sometimes compare e-assessment practices between courses (e.g., the course we targeted provided delayed feedback while another course provided immediate feedback during quiz attempts). Further study in other contexts would therefore be welcome, to gain insight into the role of context in shaping students' views.

In future work, we plan to use our findings to develop and validate a survey, to examine students' views on a larger scale. In their review of previous research on students' perceptions of feedback, Van der Kleij and Lipnevich (2020) found that survey methods were widespread, although "the overall rigour of survey methodology in the selected studies was poor" (p. 365), so there is scope to make a worthwhile contribution by developing and validating the survey with a rigorous approach (cf. Muis et al., 2014).

## References

Alarfaj, M., & Sangwin, C. (2022). Updating STACK Potential Response Trees Based on Separated Concerns. *International Journal of Emerging Technologies in Learning (iJET)*, *17*(23), Article 23. https://doi.org/10.3991/ijet.v17i23.35929

Bickerton, R. T., & Sangwin, C. J. (2021). Practical online assessment of mathematical proof. *International Journal of Mathematical Education in Science and Technology*. https://doi.org/10.1080/0020739X.2021.1896813

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Dorko, A. (2020). Red X's and green checks: A model of how students engage with online homework. *International Journal of Research in Undergraduate Mathematics Education*, *6*(3), 446–474. https://doi.org/10.1007/s40753-020-00113-w

Iannone, P., & Simpson, A. (2015). Students' preferences in undergraduate mathematics assessment. *Studies in Higher Education*, *40*(6), 1046–1067. https://doi.org/10.1080/03075079.2013.858683

Kinnear, G., Jones, I., Sangwin, C., Alarfaj, M., Davies, B., Fearn, S., Foster, C., Heck, A., Henderson, K., Hunt, T., Iannone, P., Kontorovich, I., Larson, N., Lowe, T., Meyer, J. C., O'Shea, A., Rowlett, P., Sikurajapathi, I., & Wong, T. (2022). A collaboratively-derived research agenda for E-assessment in undergraduate mathematics. *International Journal of Research in Undergraduate Mathematics Education*. https://doi.org/10.1007/s40753-022-00189-6

Kinnear, G., Wood, A. K., & Gratwick, R. (2022). Designing and evaluating an online course to support transition to university mathematics. *International Journal of Mathematical Education in Science and Technology*, *53*(1), 11–34. https://doi.org/10.1080/0020739X.2021.1962554

Lipnevich, A. A., & Smith, J. K. (2022). Student – Feedback Interaction Model: Revised. *Studies in Educational Evaluation*, *75*, 101208. https://doi.org/10.1016/j.stueduc.2022.101208

Muis, K. R., Duffy, M. C., Trevors, G., Ranellucci, J., & Foy, M. (2014). What were they thinking? Using cognitive interviewing to examine the validity of self-reported epistemic beliefs. *International Education Research*, *2*(1), 17–32.

Rønning, F. (2017). Influence of computer-aided assessment on ways of working with mathematics. *Teaching Mathematics and Its Applications: An International Journal of the IMA*, *36*(2), 94–107. https://doi.org/10.1093/teamat/hrx001

Van der Kleij, F. M., & Lipnevich, A. A. (2020). Student perceptions of assessment feedback: A critical scoping review and call for research. *Educational Assessment, Evaluation and Accountability*, *33*(2), 345–373. https://doi.org/10.1007/s11092-020-09331-x

# INVESTIGATING STUDENTS' UNDERSTANDING OF ALGEBRAIC LETTERS USING LATENT CLASS ANALYSIS

Katrin Klingbeil[1] and Filip Moons[2]

[1]University of Duisburg-Essen, Germany; [2]Freudenthal Institute, Utrecht University, The Netherlands

*To design valid assessment tools, it is necessary to understand what hurdles, common errors and misconceptions students encounter in the tested domain. Identifying typical patterns of thinking can be helpful to diagnose and communicate students' understanding to teachers. In this report, we investigate response patterns of 2051 German Year 7 and 8 students to six multiple-choice tasks of the SMART test "Meaning of Letters" that has been designed to assess the letter-as-object misconception. Using Latent Class Analysis, six response patterns could be identified. These patterns are described and analysed, and implications for improving the current assessment discussed.*

## INTRODUCTION

Assessing students' (mis)conceptions is a challenging task. SMART ("Specific Mathematics Assessments that Reveal Thinking") online tests, developed at the University of Melbourne since 2008, offer a solution by facilitating easy provision and processing of diagnostic tasks on students' conceptual understanding and potential misconceptions. SMART's extended analysis detects patterns between diagnostic tasks, revealing insights into students' understanding and misconceptions. In addition to this automatic diagnosis, it also provides teachers with explanations, tasks, and suggestions for targeted interventions (Steinle et al., 2009).

The here investigated test *Meaning of Letters* aims to assess the *letter-as-object* misconception and additionally indicates whether they show a certain subtype of this misconception based on students' responses to six multiple-choice (MC) items. Despite known challenges of MC items, developers argue that well-designed items can effectively unveil students' thinking: Klingbeil et al. (accepted) showed that students' explanations aligned well with their shown (mis)understandings in their MC responses. The SMART diagnosis in form of three stages of understanding and the flagging of the subtype is intended as a helpful simplification to support teachers in reacting to students' needs timely, concrete and without too much effort. However, this simplified allocation seems not sufficient for a detailed scientific analysis of students' understanding and the development thereof. This is especially the case since in our sample, only 21 of 2051 students were diagnosed at the highest stage. Therefore, we have conducted a latent class analysis to identify more specific response patterns concerning the *letter-as-object* misconception.



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 145-152). PME.

## THEORETICAL BACKGROUND

Arcavi, Drijvers and Stacey "distinguish five facets of the concept of variable: a placeholder for a number, an unknown number, a varying quantity, a generalised number, and a parameter" (2017, p. 12). Across these facets, variables stand for or refer to one or more numerical values. Yet, algebra learners often struggle with this numerical interpretation, and various typical errors and misconceptions have been identified. One of them, the *letter-as-object* misconception, has been described by Küchemann in 1981 as the letter being "regarded as a shorthand for an object or as an object in its own right" (p. 104) and extensively documented over decades (e.g., Akhtar & Steinle, 2017). As part of the foundational CSMS study on the mathematical understanding of secondary school students in the United Kingdom, Küchemann (1981) utilised the following task:

> "Blue pencils cost 5 pence each, and red pencils cost 6 pence each. I buy some blue and some red pencils and altogether it costs me 90 pence. If $b$ is the number of blue pencils bought and if $r$ is the number of red pencils bought, what can you write down about $b$ and $r$?" (p. 107)

While only 10% of tested 14-year-old students provided the correct equation $5b + 6r = 90$, 17% gave $b + r = 90$ as an answer, which might have been read as "blue pencils and red pencils together cost 90 pence" (LO). Interpreting b as "the number of blue pencils" is a possibility here, too; however, this would still imply a wrong understanding of equations with a number of pencils on one side of the equation and the price of the pencils on the other. Interestingly, 6% of the students came up with another kind of equation: $6b + 10r = 90$ or $12b + 5r = 90$. These students had figured out a possible solution to the problem first and then used these values as coefficients in their equation. Since the letters are used as abbreviations for the involved objects ("12 blue pencils and 5 red pencils together cost 90 pence"), this is regarded as a special form of LO, which we will refer to as the *solution-as-coefficient* (SAC) misconception in the following. Another special form of LO is called *letter-as-unit* (LU) when the algebraic letter is interpreted as an abbreviation for a unit (Akhtar & Steinle, 2017), e.g., in a task about 8 trucks weighing 24 tonnes, the $t$ in the equation $8t = 24$ would be misinterpreted as standing for tonnes (not realising that this would not be a correct equality).

### Research question

The *Meaning of Letters* SMART test assesses the understanding of variables and detects the presence of the *letter-as-object* misconception, leading to the following research question: Which response patterns regarding the *letter-as-object* misconception can be identified among German grade 7 and 8 students based on their responses to the six multiple-choice tasks of the SMART test *Meaning of Letters*?

## METHODS AND MATERIALS

### SMART test *Meaning of Letters*

For the diagnosis of students, two parallel versions of the SMART test *Meaning of Letters* were used with the A or B version randomised by class. Here, we describe only the A version of the German translation of the test (see Figure 1).

The first task type (Meaning tasks), originating from the work of MacGregor and Stacey (1997), uses only one algebraic letter and asks students to decide on the meaning of this letter in a linear equation in a given context. While the *Ducks* item uses the initial letter of the involved objects and units, the *Bricks* item uses the letter $y$. Apart from the correct response (cost/height), MC options include the involved objects (singular and plural; LO) as well as the corresponding unit (LU).

| **Meaning tasks** | **Additive tasks** | **Proportional tasks** |
|---|---|---|
| *Ducks* | *Garden* | *Biros* |
| Lucy bought 6 ducks for \$12. She wrote the equation $6d = 12$. What does $d$ in Lucy's equation stand for?<br>☐ the cost of one duck (COR)<br>☐ one duck (LO)<br>☐ ducks (LO)<br>☐ dollars (LU) | Payam bought $r$ red rose bushes and $l$ lilac lavender bushes. The roses cost €4 each. The lavenders cost €5 each. Which equation says that the total cost was \$70?<br>☐ $4r + 5l = 70$ (COR)<br>☐ $r + l = 70$ (LO)<br>☐ $10r + 6l = 70$ (SAC) | Biros are sold in packs of 3. Sam bought $p$ packs and got $b$ biros altogether.<br>Choose the correct equation:<br>☐ $3p = b$ (COR)<br>☐ $p = 3b$ (LO)<br>☐ $p = 3$ (LO)<br>☐ $b + p = 4$ (LO)<br>☐ $30b = 10p$ (SAC) |
| *Bricks* | *Wheels* | *Racetrack* |
| Tina stacked 9 identically sized bricks on top of each other making a 99 mm high tower. She wrote the equation $9y = 99$. What does $y$ in Tina's equation stand for?<br>☐ the height of one brick (COR)<br>☐ one brick (LO)<br>☐ the bricks in the tower (LO)<br>☐ millimetres (LU) | At a bike shop there are $b$ bikes (2 wheels) and $t$ trikes (3 wheels). Which equation says that there is a total of 100 wheels in the shop?<br>☐ $2b + 3t = 100$ (COR)<br>☐ $b + t = 100$ (LO)<br>☐ $35b + 10t = 100$ (SAC) | A car takes 12 minutes to drive round this racetrack. A driver drives $r$ times around the racetrack in $m$ minutes. Choose the correct equation:<br>☐ $12r = m$ (COR)<br>☐ $12m = r$ (LO)<br>☐ $r = 12$ (LO)<br>☐ $5r = 60m$ (SAC) |

**Figure 1: Tasks of *Meaning of Letters* test**

The second type of task (Additive tasks) is based on Küchemann (1981). It uses two algebraic letters (corresponding to the initial letters of involved objects), which are additively connected and restricted by the given situation. Students are supposed to choose the correct linear equation (in standard form) for the described context. In the correct equation, the letters represent the number of objects and the coefficients for the price per object (*Garden*) and the number of components per object (*Wheels*), respectively. The first alternative response option simply adds the variables without any coefficients, making it possible to interpret the letters as abbreviations for the involved objects (e.g., "Bikes and trikes have 100 wheels altogether."; LO). In the equation of the other alternative option, coefficients equal a possible solution to the problem (that has not been posed) so that the equation can be read as some solution sentence (e.g., "35 bikes (with 2 wheels each) plus 10 trikes (with 3 wheels each) have 100 wheels altogether."; SAC). Also, in this case, the letters are read as abbreviations for the involved objects.

The third task type (Proportional tasks) is derived from the famous "Students and Professors" problem (Clement et al., 1981):

> "There are six times as many students as professors at this university." Write an equation using S for the number of students and P for the number of professors.

This is often answered with $6S = P$ instead of $6P = S$. The proportional tasks in the SMART test have the same algebraic structure: the two variables are directly proportional to each other. Students are again asked to choose the equation matching the given situation. In the correct equation, the letters (matching the initial letters of involved objects/units) stand for the number of objects for both involved objects (*Biros*) or for the number of racetrack rounds and the number of minutes (*Racetrack*). In both items, the coefficient is the proportionality constant (number of biros per pack or number of minutes per round). The first alternative response option is the reverse of the correct equation, which allows for a LO interpretation (e.g., "A pack contains 3 biros." or "1 round equals 12 minutes"; LO). For the *Racetrack* task, the first alternative can also be seen as an LU interpretation (e.g., "12 minutes equals one round). However, since it is unclear how exactly students interpret the letter here, we opted for the more general LO interpretation. The LO interpretation also applies to the second alternative response although the second variable is missing (e.g., as "A pack has 3."; LO). In the equation of the third alternative option, the coefficients correspond to a possible solution (to the question that has not been asked), which can be interpreted as a kind of solution sentence (e.g., "Sam bought 10 packs and has 30 biros now."), indicating the SAC misconception. The *Biros* task offers one more response option that features the addition of the two variables without coefficients. Again, the letters can be interpreted as abbreviations (e.g., as "One biro plus a pack of biros is 4 altogether.", LO). This response type does not make sense for the *Racetrack* task since no different objects but rounds and minutes would be added.

Students with less than 2 correct tasks are allocated at Stage 0, students with 2 to 5 correct tasks at Stage 1, and students with 6 correct tasks at Stage 2. If at least one of the SAC responses has been chosen, SAC will be flagged for this student.

**Participants**

2051 students from six federal states of Germany completed the test. These students were taught by 103 mathematics teachers, leading to a nested data structure. The students were in grade 7 or 8 (aged 12–14), and the SMART online test was taken 1–2 weeks into a teaching sequence about variables, algebraic expressions and/or equations. Thus, in grade 7, students should have been familiar with a basic concept of variables and be able to use and manipulate them in easy algebraic expression before taking the test. In grade 8, students probably have started focussing on (solving) equations.

The overall response rates are shown in Figure 2. Note the low number of correct answers and answers with the LU misconception. The LO misconception was omnipresent in the responses to meaning and proportional tasks while SAC was present in most responses to additive tasks.

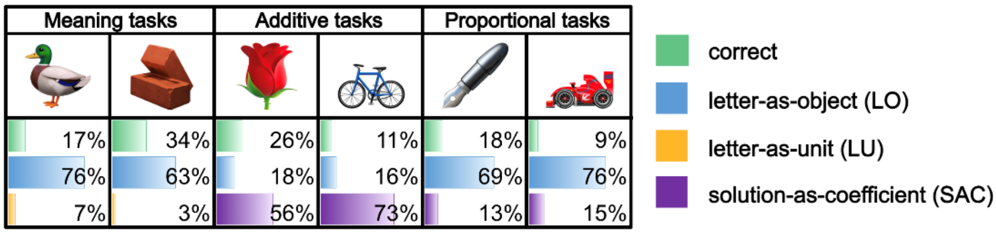



**Figure 2: Overall response rates to the *Meaning of Letters* test**

### Data analysis

The response patterns of the test were analysed using Latent Class Analysis (LCA) (Brandenburger, & Schwichow, 2023). LCA is a form of structural equation modelling useful for identifying patterns/groups within categorical responses. These patterns/groups are called *latent classes*. LCA considerably reduces the complexity of the data by grouping students with similar patterns of responses together in one class, bringing down the 2051 individual response patterns to a comprehensible, clearly distinguishable number of latent classes. We used SAS Enterprise Guide 8.3 with the PROC LCA for the LCA analysis, considering the nested data structure. As students were not required to answer all questions, 49 students of the 2051 (2.3%) left some tasks unanswered. Hot-deck imputation was used to impute these missing values.

## RESULTS

### Model selection and model fit

We iteratively built several models with different numbers of classes to choose the appropriate number of classes for our LCA model. Information criteria (AIC/BIC) and the possibility of giving meaningful labels to the classes were used to decide the number of classes. The 6-class model had the best information criteria, the most interpretable latent classes, and no particularly infrequent class. Hit rates (mean probabilities for the best fitting class) were between 70% and 91% indicating a reasonably good reliability.

### Description of the latent classes

In Figure 3, the six classes with their prevalence and their item-response probabilities for each task are shown.

Class 1 (prevalence: 23%) is characterised by LO responses being most likely in all tasks. Even when the initial letter of the involved object is not used (*Bricks*), LO is more likely than a correct answer. The subtype SAC is also possible in additive tasks, but less likely than LO. This indicates a relatively consistent interpretation of letters as abbreviations for involved objects. We label this class "*LO predominant*".

Class 2 (prevalence: 11%) is characterised by high probabilities for SAC responses. However, for *Biros* a LO response is more likely than a SAC response. Since this is the class with the highest probability for SAC in proportional tasks, students in this class seem to be quite convinced that coefficients stand for (possible) solutions also in different equation types. We label this class "LO with SAC".

| Class | Meaning tasks | | Additive tasks | | Proportional tasks | | SMART diagnosis | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | 0 | 0 + SAC | 1 | 1 + SAC | 2 |
| 1. LO predominant | 12% | 25% | 8% | 10% | 19% | 11% | | | | | |
| (Prevalence: 23%) | 81% | 70% | 64% | 57% | 71% | 76% | 152 | 191 | 63 | 37 | 0 |
| | 7% | 4% | 28% | 33% | 10% | 13% | | | | | |
| 2. LO with SAC | 8% | 33% | 15% | 0% | 8% | 6% | | | | | |
| (Prevalence: 11%) | 88% | 66% | 4% | 1% | 54% | 0% | 0 | 13 | 0 | 38 | 0 |
| | 4% | 1% | 81% | 99% | 37% | 94% | | | | | |
| 3. LO with SAC only in additive items | 3% | 25% | 8% | 2% | 9% | 2% | | | | | |
| | 92% | 73% | 2% | 0% | 79% | 98% | 3 | 695 | 1 | 64 | 0 |
| (Prevalence: 38%) | 5% | 2% | 90% | 98% | 12% | 0% | | | | | |
| 4. Correct meaning items with | 89% | 82% | 23% | 7% | 11% | 2% | | | | | |
| LO/SAC elsewhere | 0% | 13% | 3% | 0% | 79% | 85% | 0 | 19 | 13 | 130 | 0 |
| (Prevalence: 8%) | 11% | 5% | 74% | 93% | 10% | 13% | | | | | |
| 5. LO apart from additive items | 16% | 24% | 91% | 39% | 29% | 8% | | | | | |
| (Prevalence: 15%) | 75% | 70% | 9% | 17% | 64% | 89% | 23 | 61 | 221 | 57 | 0 |
| | 9% | 6% | 0% | 44% | 7% | 3% | | | | | |
| 6. Mostly correct | 61% | 86% | 69% | 31% | 85% | 64% | | | | | |
| (Prevalence: 5%) | 30% | 12% | 10% | 7% | 14% | 33% | 1 | 0 | 48 | 30 | 21 |
| | 9% | 2% | 21% | 62% | 1% | 3% | | | | | |

**Figure 3: Latent classes with their item-response probabilities for every task and SMART diagnosis for students with this as best fitting class**

Class 3 (prevalence: 38%) is characterised by very high probabilities for SAC responses in additive tasks and high probabilities for LO responses in all other tasks. We label this class "LO with SAC only in additive tasks".

Class 4 (prevalence: 8%) is characterised by high probabilities for correct responses in meaning tasks, high probabilities for SAC in additive tasks, and high probabilities for LO in proportional tasks. Students in this class seem to be able to identify the correct meaning of an algebraic letter when directly being asked for it. However, when choosing equations, they still fall into the trap of interpreting letters as abbreviations. We label this class "*Correct meaning tasks with LO/SAC elsewhere*".

Class 5 (prevalence: 15%) is characterised by medium to high probabilities for LO in all tasks other than additive tasks. While in *Garden* the correct response is most likely, in *Wheels* SAC is slightly more likely than the correct response. We label this class "*LO apart from additive tasks*".

Class 6 (prevalence: 5%) is characterised by correct responses being most likely in almost all tasks. Only in *Wheels*, SAC is double as likely as the correct response. In *Ducks* and *Racetrack*, LO is also possible but half as likely as the correct response. This indicates at least a partial understanding that algebraic letters do not stand for abbreviations. We call this class "*Mostly correct*".

Comparing students' best fitting classes with their SMART diagnosis, in general, a good concordance can be seen. For example, students with Class 2, 3 or 4 as best fitting class were also mainly diagnosed as having SAC. Deviations can be explained by the fact that students' response patterns do not necessarily perfectly match their best fitting class. Furthermore, it can be seen that being diagnosed at Stage 1, for example, does not allow for any inferences regarding the best fitting class.

## Discussion and Outlook

Utilising LCA, six distinct response pattern classes were identified, offering detailed insights into the relationship between students' comprehension, misconceptions, and test tasks. These classes play a crucial role in enhancing our understanding of how students interpret algebraic letters across various contexts. It is important to note that a comprehensive understanding of the implications is an ongoing research process, and this discussion marks our initial attempt at exploring these insights.

Starting with classes that exhibit at least some correct answers, a notable discovery is that Class 6, characterised by mostly correct answers, has a low prevalence of 5% and still shows many SAC responses to the *Wheels* task. The absence of a class labelled 'All answers correct' is not surprising, as only 21 students (1%) would belong to this class, which contradicts the principle of a good LCA model that avoids very rare classes. Class 4 (8%) is intriguing, displaying high probabilities for correct meaning tasks, but struggles when translating this understanding into equations. These students seem to possess a superficial knowledge of variable meanings, adequate for direct inquiries about meaning with one variable but insufficient when dealing with equations involving two variables. This underscores the importance of recognising that merely asking about the meaning of letters in simple contexts does not necessarily imply a deep and accurate understanding. Class 5 (15%) is characterised by a very high probability of a correct answer on the *Garden* task and LO/SAC in most other tasks. Since these students do not seem to grasp the meaning of letters, it is likely that these correct responses are not a result of (partial) understanding but of a strategy of combining given letters and numbers according to the described situation without proper understanding.

Regarding the LO misconception and its subtypes, it is crucial to highlight that the LU misconception had a minimal occurrence in the meaning tasks. Some students consistently show LO (Class 1, 23%); however, even in this class, the subtype SAC has a probability of 33% in *Wheels*. This might indicate that this task especially fosters students' urge to come up with a numerical solution. In general, the subtype SAC is often clearly present in additive tasks only (especially Classes 3 and 4). Probably, such additive equations can be read more intuitively as a solution sentence such as "35 bikes plus 10 trikes have 100 wheels altogether" compared to proportional tasks that would have to be read as something like "Sam bought 30 biros in 10 packs". It is also possible that the additive tasks more easily trigger some students' desire to give a solution than proportional tasks. In this regard, Class 2 is exceptional: a very high probability for SAC in *Racetrack,* but only 37% in the *Biros* tasks, while they are both proportional tasks (with similar response rates, see Figure 2). This might indicate that a SAC interpretation in proportional equations is more likely when the letters involved refer to non-physical objects (e.g., rounds in *Racetrack)* or can be interpreted as units (e.g., minutes in *Racetrack).*

The analysis of the identified classes underlines how important the task type, complexity, and context – including the realness of involved entities and the underlying structure of the equation – seems to be for correctly interpreting algebraic letters. These are aspects that need to be considered for teaching as well as assessment. However, for a quick assessment that teachers can immediately react to, the information of the identified classes, on the one hand, might be too detailed, on the other hand, the informative "all answers correct" class (corresponding to Stage 2) is missing. Therefore, it needs to be examined further in what way the identified classes may improve the *Meaning of Letters* test.

## REFERENCES

Akhtar, Z., & Steinle, V. (2017). The prevalence of the 'letter as object' misconception in junior secondary students. In A. Downton, S. Livy & J. Hall (Eds.), *Proceedings of the 40th annual conference of the Mathematics Education Research Group of Australasia* (pp. 77–84). MERGA.

Arcavi, A., Drijvers, P., & Stacey, K. (2017). *The learning and teaching of algebra: Ideas, insights, and activities*. Routledge.

Brandenburger, M., Schwichow, M. (2023). Utilizing Latent Class Analysis (LCA) to Analyze Response Patterns in Categorical Data. In X. Liu, & W. J. Boone (Eds.), *Advances in Applications of Rasch Measurement in Science Education. Contemporary Trends and Issues in Science Education*, (Vol. 57, pp. 123–154). Springer, Cham. https://doi.org/10.1007/978-3-031-28776-3_6

Clement, J., Lockhead, J., & Monk, G. (1981). Translation difficulties in learning mathematics. *American Mathematical Monthly*, *88*(4), 286–290.

Klingbeil, K., Rösken, F., Barzel, B., Schacht, F., Stacey, K., Steinle, V. & Thurm, D. (accepted). Validity of multiple-choice digital formative assessment for assessing students' (mis)conceptions: Evidence from a mixed-methods study in algebra. In H.-G. Weigand, J. Trgalova & M. Tabach. Mathematics Teaching, Learning and Assessment in the Digital Age. *ZDM – Mathematics Education, 56* (4).

Küchemann, D. (1981). Algebra. In K. M. Hart, M. L. Brown, D. E. Küchemann, D. Kerslake, G. Ruddock, & M. McCartney (Eds.), *Children's understanding of mathematics: 11–16* (pp. 102–119). John Murray.

MacGregor, M., & Stacey, K. (1997) Students' understanding of algebraic notation: 11-16. *Educational Studies in Mathematics*, *33*(1), 1–19.

Steinle, V., Gvozdenko, E., Price, B., Stacey, K., & Pierce, R. (2009). Investigating students' numerical misconceptions in algebra. In R. Hunter, B. Bicknell & T. Burgess (Eds.), *Proceedings of the 32nd annual conference of the Mathematics Education Research Group of Australasia* (Vol. 2, pp. 491–498). MERGA.

# DOES STUDENTS' CREATION OF PROBLEM-SOLVING VIDEOS ENHANCE THEIR EXAM PERFORMANCE?

Igor' Kontorovich[1], Hongjia H. Chen[1], Ian Jones[2],
Nicolette Rattenbury[1], and Padraic Bartlett[1]

[1]The University of Auckland, [2]Loughborough University

*We explore the impact of undergraduates' creation and peer reviewing of problem-solving videos on their exam performance. In a large first-year course for non-mathematics majors, students were provided with a bank of problems from past exams with historically low scores. As part of the homework assignment, the students video-recorded a solution to a problem of their choice, while elaborating on the involved concepts and steps. Then, the submitted videos were randomly allocated for peer reviewing. We consider this activity through the lens of effective digital task design and deep active learning. Quantitative results offer some evidence of the positive impact of the activity on students' performance on similar problems in a final exam.*

## RATIONALE, EXTANT RESEARCH, AND AIM

The calls for active learning in undergraduate studies sound louder than ever (e.g., Freeman et al., 2014). Active learning is a broad construct, encompassing a wide range of pedagogies that include "all kinds of learning beyond the mere one-way transmission of knowledge in lecture-style classes […]. It requires engagement in activities (writing, discussion, and presentation) and externalizing cognitive processes in the activities" (Mizokami, 2018, p. 79). However, incorporating active-learning activities is far from trivial in undergraduate mathematics (e.g., Conference Board of the Mathematical Sciences, 2016). This is especially challenging in first-year courses, which have traditionally been content-dense, lecture-based, and taken by large student cohorts from diverse mathematics backgrounds. These challenges draw attention to active learning initiatives that were successfully integrated into undergraduate courses. We explore one such initiative in this study.

Many students create video content on a regular basis, which highlights the possibility of leveraging this modern way of digital existence for the benefit of university studies. Educational research has been exploring activities where students create videos around the course content and peer-review each other's creations (e.g., Huang et al., 2020; Lazzari, 2009). Hulsizer (2016) refers to student-generated videos as "an active, deliberate, and cooperative exercise" (p. 271). The engagement with the peers' videos provides additional opportunities to interact with the content, use mathematical terminology, apply critical thinking, and practice feedback provision. Contemporary learning management systems (e.g., Canvas, Moodle) can support the streamlining of submission and peer reviewing of student-generated videos even in large courses.

Only a handful of studies have explored the use of student-created videos in the context of university mathematics (e.g., Hulsizer, 2016). These studies report on the



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 153-160). PME.

advantages of videos and knowledge acquisition as these were perceived by the students. Yet, research into the learning gains of these activities remains scarce (e.g., Huang et al., 2020). Indeed, Hulsizer (2016) notes that "it is the hope that video creation will promote retention of knowledge and a deeper understanding of the material" (p. 271). To understand whether there might be evidence for this hope, this study aims *to explore the impact of undergraduates' creating and peer-reviewing mathematical videos on their learning*.

## THEORETICAL BACKGROUND

### Deep active learning

Matsushita (2018) notices that active learning agendas often focus on learning formats and argues that "learning in universities ought to be not only active but also deep" (p. 15). Drawing on Engeström theory, Matsushita introduces the notion of *deep active learning* and associates it with a six-step learning cycle: (i) *motivation* that emerges from a conflict between problems that students encounter and their existing state of knowledge; (ii) *orientation* that pertains to student engagement in activities to resolve the conflict; (iii) *internalization* or knowledge acquisition; (iv) *externalization* of the acquired knowledge through resolving the original conflict; (v) *critique* that can occur from students realizing the limits of their previous or acquired knowledge and the new need to reconstruct it; and (iv) *control*—looking back at the carried out sequence of processes and revising them before moving on. Matsushita emphasizes that the cycle is applicable to different time spans, including short-term multi-stage activities.

### Effective task design in a blended environment

Albano et al. (2021) introduce a framework for effective task design in a blended environment. The framework is a tetrahedron model for didactic systems in e-learning. The four vertices of the tetrahedron are: *mathematics* to be taught and studied, a s*tudent* who is expected to learn it, a *tutor* who supports the student's learning, and a *designer* who is in charge of developing and implementing the activity. Technology features twice in the model: as an inscribed entity, representing the set of digital resources and tools the actors utilize for explicit didactic purposes; and external technology that the actors use on a regular basis and that can be involved in the particular case as well.

Two comments are in order. First, the four vertices do not denote fixed entities or people, but roles that can be played by different actors. For example, a technological resource can act as a tutor. Second, Albano et al. (2021) argue that the designer's role

> is best played by a collective that may include mathematics teachers, experts in mathematics education and in the specific mathematical subject at stake and experts in the use of technology. The "collective" evokes a co-disciplinary perspective for the design: it allows the co-thinking, the integration and the negotiation of the right balance between the various viewpoints […] and expertise in the collective focused on the same object of study.

To design a specific task, different facets of the tetrahedron are analyzed with attention to *cognitive*, *epistemological*, and *didactic* dimensions. The cognitive dimension focuses on the learning difficulties students encounter when studying a particular mathematical topic. These include common misconceptions and external factors, such as previous experiences with mathematics. The *epistemological* dimension revolves around learning challenges that stem from the nature of the focal piece of mathematics. The didactic dimension encompasses the organizational aspects of the course of teaching (e.g., assessment).

## THE STUDY

This is a field-driven initiative conducted in collaboration between university teachers and educational researchers. It is a research-infused teacher-led innovation, where the collaborating partners mobilize their professional knowledge to execute a disciplined inquiry in an authentic learning setting (Kontorovich et al., 2023). This collaborative aspect is consistent with the role of a designer in Albano et al.'s (2021) framework.

### Setting, video task, and a priory analysis

Our data came from a *video task* the teachers developed as part of their instruction of a large first-year course for non-mathematics majors, typically ranging between 150 and 700 students. The course covered standard topics in Calculus and Linear Algebra. Consistently with the cognitive dimension of Albano et al.'s (2021) model, the teachers constructed a *bank* of 30 problems from past exams with especially low success rates. An epistemological analysis was conducted to delineate potential sources of challenges that the concepts involved in the problems entailed (e.g., Kontorovich, 2022).

In the video task, students were asked to select a problem from the bank and record a 1–5 minute-long video of its solution. The guidelines specified that the solution was expected to contain explanations for the executed methods as if the students were communicating with a peer who was having a "hard time" with the course concepts.

The problems in the final exam (and, accordingly, in the bank) were multiple-choice. The students were given around ten days to submit their solutions via Canvas. Then, each solution was randomly assigned to four peers, who reviewed it independently in terms of mathematical validity and clarity. The final answers to the bank problems were published before peer reviewing. The task was given in the last week of the semester, followed by a final exam within a few days.

Figure 1 offers a tetrahedron representation of the video task (Albano et al., 2021). It can be characterized through three main phases: video-recording, peer-reviewing, and feedback-interiorization. Each phase not only possessed the attributes of active learning (e.g., communication, presentation, externalizing thinking processes), but also made room for a deep learning cycle (Matsushita, 2018). The bank problems are in the focus of the first phase. They were selected based on the unsatisfactory performance of previous course cohorts in final exams. Thus, it was reasonable to propose that the study students may also have experienced challenges with these problems (motivation in terms of Matsushita, 2018). The video task asked students to resolve the challenge by solving a problem and producing a peer-oriented video of the solution (orientation and externalization). These activities involve internalization of new knowledge and revisiting of the existing one to develop a new technology-based resource for the course (an inscribed technology in terms of Albano et al., 2021).

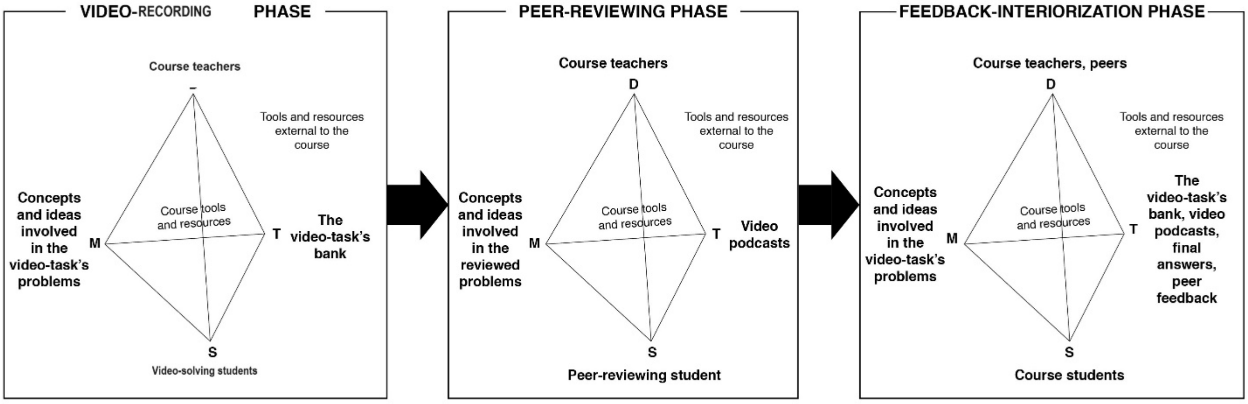



Figure 1: Tetrahedron model of the video task.

The video task encompasses multiple opportunities for critique and control. First, it is assigned on the completion of the course instruction and is presented to students as a productive means for course review. Accordingly, the students can use their work on the bank problems to get a sense of their grasp of the content, monitor the limits of their knowledge, and revisit relevant topics. Second, the students are given a choice regarding the problem to video-solve. This summons engagement with several problems in the bank to make an aware choice. Third, as reviewers, the students are expected to critique and control their peers' videos. Critique and control are especially needed when the reviewed problems are different from the one a reviewer chose to solve in the previous phase and when the peers' solution is different from the one a reviewer would execute. This is also the stage where the produced video solutions can act as tutors (Albano et al., 2021). Fourth, final answers to bank problems and peers' comments on the submitted solution provide additional opportunities to look back at the conducted work, check its validity, and revise the course content. To summarize, the video task encompassed multiple opportunities for deep active learning.

## Analysis

The questions underpinning the analysis were: (i) Does students' performance on a certain problem in the exam depend on its similarity to the problems they video-solved and/or peer-reviewed in the video task? In other words, do students, who submit solutions to and/or peer review problems that are similar to the ones in the final course exam, perform better, compared to students who submitted a video solution to a different problem?; and (ii) Did exam performance change after the introduction of the video task? Note that both questions concern students' performance on specific problems in the exam that were similar to those in the bank of the video task.

We operate with data from three semesters when students sat the exams in closed-book invigilated conditions. The analysis started with the identification of 16 problems from the focal exams that were sufficiently similar to those in the bank in terms of the involved mathematical concepts, question formulations, and options for an answer. For example, we matched the bank problem "What is the area between the curve $x^2 - 1$ and the $x$-axis over the interval $[0,2]$? (a) 6; (b) $\frac{5}{3}$; (c) 2; (d) $\frac{2}{3}$" to the exam problem, asking "What is the area between the graph of the function $f(x) = x^2 + x - 2$ and the $x$-axis over the interval $[0,4]$? (a) $\frac{71}{3}$; (b) $\frac{64}{3}$; (c) $-\frac{64}{3}$; (d) 0."

To pursue the first research question, for each identified problem, we split the student cohort into two groups: (a) students who engaged with matching bank problems as solvers, peer reviewers, or both; and (b) those who worked on other problems from the bank. Then, we used Pearson's Chi-squared test with Yates' continuity correction to compare the performance between the groups.

Regarding the second question, we used historical data on exam performance in the last decade. While the course underwent several changes over the years, its syllabus, structure, and instruction remained relatively stable. This enabled us to compare performance before and after the introduction of the video task. We identified 21 problems from past exams that were sufficiently similar to the exam problems in the studied semesters. Then, we compared relative exam performance of the historical and focal cohorts on the corresponding sub-parts of the exams (i.e., Calculus or Linear Algebra). If the cohorts performed similarly on the whole topic, the comparison in performance on a certain historical-focal dyad may be indicative of the impact of the video task. Specifically, if the task entailed deep active learning, it could be expected to show in dyads that had a similar problem in the bank. In turn, the performance on dyads that did not have a close "relative" in the bank could be expected to be similar.

## RESULTS

Overall, students who engaged with the bank problems that were matched with final exam problems outperformed students who engaged with the bank problems that were unmatched with exam problems. More precisely, the matched-problem students scored more highly than the unmatched problem students on 14 out of 16 problems, and this was significant in 5 cases (Table 1). In other words, the unmatched problems group scored more highly only on two problems, but the differences were not significant.

We further broke down the abovementioned results in terms of the number of activities undertaken by students (recall that some students created as well as reviewed videos of the same matched problem, and some students reviewed multiple videos of a given problem). For each problem, the number of students for a given activity varied from 1 to 178, making statistical testing inappropriate. Yet, this breakdown allowed a sense of a pattern: for "easy" problems correctly solved by more than 75% of the students, there was no overall systematic advantage of engaging in video creation over video reviewing or vice versa, or even of engaging in both activities, or of engaging in reviewing multiple times. For more "difficult" problems correctly solved by less than 50% of the students, the creation of video solutions was more advantageous than peer-reviewing. We should be cautious because the highest percent correct exam problems are mostly for activities with which only a few students engaged.

| ID | 1 | 2 | 3 | 4 | 5 | 6 | 7** | 8 |
|---|---|---|---|---|---|---|---|---|
| Match | 25.3% (170) | 26.9% (104) | 27.0% (281) | 30.0% (40) | 66.5% (355) | 68.5% (54) | 68.7% (339) | 68.9% (106) |
| No match | 26.6% (587) | 26.2% (653) | 25.8% (476) | 26.1% (717) | 55.3% (237) | 72.3% (47) | 58.1% (418) | 59.8% (651) |
| ID | 9*** | 10* | 11 | 12 | 13 | 14 | 15*** | 16*** |
| Match | 72.9% (280) | 73.3% (45) | 74.7% (170) | 76.0% (25) | 84.1% (339) | 89.7% (58) | 93.0% (355) | 93.2% (354) |
| No match | 56.1% (312) | 55.0% (712) | 71.9% (587) | 70.6% (567) | 82.5% (418) | 83.7% (43) | 84.0% (237) | 83.6% (238) |

Table 1: Percentage correct (number of students) for exam problems with/without a matched problem. Significance: .05*, .01**, .001*** (Holm-Bonferroni correction).

| | Focal cohort | | | | Historical cohort | | |
|---|---|---|---|---|---|---|---|
| Problem dyad | Matching | | Not matching | | | | |
| | % | $n$ | % | $n$ | % | $n$ | $p$ |
| 1 | 73.3 | 45 | 54.9 | 712 | 39.3 | 878 | <.001* |
| 2 | 68.9 | 106 | 59.8 | 651 | 69.8 | 149 | .983 |
| 3 | 68.9 | 106 | 59.8 | 651 | 20.9 | 878 | <.001* |

Table 2: Student performance on matching problems before and after the introduction of the video task (with a similar problem in the bank)

Table 2 presents students' performance on matching problems in final exams before and after the introduction of the video task for three dyads that had a similar problem in the bank. For dyad 1, the lowest-scoring (non-matching) focal cohort outperformed the past cohort and this difference was significant, $\chi^2(1, N = 1{,}590) = 37.97$, $p < .001$. The same held for dyad 3, $\chi^2(1, N = 1{,}529) = 238.49$, $p < .001$. For dyad 2, the past cohort outperformed the highest-scoring (matching) focal cohort, but this difference was not significant, $\chi^2(1, N = 255) < .01$, $p = .983$. Then, in two of the three cases, the cohorts that undertook the video task performed significantly better than their predecessors. The results for dyad 2 are barely surprising given the past cohort was a small summer-school cohort that tends to perform better than students in regular semesters. Indeed, the same problem from the focal semester was used in dyad 2 and 3 (hence the identical numbers in the table). In 11 out of 18 dyads that did not have a similar problem in the bank, there was no significant difference in the performance of the focal and historical cohorts. Overall, this analysis supports the positive impact of the video task on students' exam performance.

## CONCLUDING REMARKS

Our analysis showed that overall, the students who created videos or peer-reviewed a solution to a certain problem, performed better on a similar problem in a final exam, compared to the students that engaged with a different problem in the bank. The design of the video task was consistent with Matsushita's (2018) theory, and thus, it appears viable to propose that the task spurred cycles of deep active learning, indeed. Note that our analysis revolved around single problems, the solutions to which the students submitted and peer-reviewed. But recall that the whole bank was accessible to all students, giving each of them an opportunity to enhance their knowledge in preparation for the final exam. Accordingly, the students who were identified as engaging with a non-matching problem could still have solved relevant problems (without making a video) and verified their work against the published answers. In these circumstances, the significant differences between the two groups may be interpreted as the added value of preparing a video-recorded solution. In a similar vein, the analysis showed that, compared to the previous cohorts who did not engage in the video task, the focal cohort performed better when there was a matching problem in the bank. More often than not, there was no significant difference between the focal and the historical cohorts when there was no matching problem in the bank.

Our results are consistent with Huang et al. (2020), who also identified a positive impact of video generation on mathematical learning and problem solving in the case of elementary students. Lazzari (2009) accounts for this impact as follows:

> designing, developing, recording and publishing lessons [or videos], experience the perspective of being listened to an evaluated first of all by peers (their colleagues) […] compels students to an extra-effort […] that leads them to a more intense and effective learning process, well beyond the simple assimilation of concepts or even their re-elaboration, up to the search for the meaning of what they are studying (p. 33).

We acknowledge that a pre-test – post-test design or including a more straightforward control group would have provided causal evidence for the impact of the video task. Such designs may be implemented by further studies. Our research studied an authentic innovation that was conceived and implemented by the course teachers to align with the needs and interests of their students. As with any research methodology, a research-infused teacher-led innovation comes with its limitations and affordances (Kontorovich et al., 2023). We believe that the key affordance of our design is offering evidence, pointing at the positive impact of the creation of videos and their peer-reviewing on student mathematical performance. The evidence emerged from a large first-year mathematics course, which shows the feasibility of the innovation in the field of practice. As Rittle-Johnson (2019) argues, "we must conduct research *within* educational settings in order to ensure the [instructional] method is feasible outside of a controlled lab setting and that the findings generalize to those settings" (p. 140).

Then, with this study, we draw attention to students' generation of mathematical videos as a modern and viable means to advance deep active learning in the tertiary context.

## ACKNOWLEDGEMENT

This work was partially supported by the British Academy [VF2\100024].

## References

Albano, G., Kondratieva, M., & Telloni, A. I. (2021). Online resources for mathematics teaching and learning at the university level: There case examples that highlight principles for task design driven by students' and teachers' needs. In A. Clark-Wilson, A. Donevska-Todorova, E. Faggiano, H.-G. Weigand, & J. Trgalová (Eds.), *Mathematics education in the digital age: Learning, practice and theory*. Routledge.

Conference Board of the Mathematical Sciences. (2016). Active learning in post-secondary mathematics education. Retrieved on 8 March 2021 from https://www.cbmsweb.org/2016/07/active-learning-in-post-secondary-mathematics-education/

Freeman, S., Eddy, S., McDonough, M., Smith, M., Okoroafor, N., Jordt, H., & Wenderoth, M. (2014). Active learning increases student performance in science, engineering, and mathematics. In *Proceedings of National Academy of Science U.S.A.*, 111(23), 8410–8415.

Huang, M. C.-L., Chou, C.-Y., Wu, Y.-T., Shih, J.-L., Yeh, C. Y. C., Lao, A. C. C., Fong, H., Lin, Y.-F., & Chan, T.-W. (2020). Interest-driven video creation for learning mathematics. Journal of Computer Education, 7(3), 395–433.

Hulsizer, H. (2016). Student-produced videos for exam review in mathematics courses. *International Journal of Research in Education and Science*, *2*(2), 271–278.

Kontorovich, I. (2023). "Find the area enclosed by …" Parceling an especially robust model of reasoning among many first-year students. *International Journal of Research in Undergraduate Mathematics Education*, *9*, 149–172.

Kontorovich, I., Chen, H., Jones, I., Rattenbury, N., & Bartlett, P. (2023). Research-infused teacher-led innovations in university mathematics education. In *Proceedings of the Thirteenth Congress of the European Society for Research in Mathematics Education* (CERME13) (pp. 2415–2422). Alfréd Rényi Institute of Mathematics and ERME.

Lazzari, M. (2009). Creative use of podcasting in higher education and its effect on competitive agency. *Computers & Education*, *52*(1), 27–34.

Mizokami, S. (2018). Deep active learning from the perspective of active learning theory. In K. Matsushita (Ed.), *Deep active learning: Toward greater depth in university education* (pp. 79–91). Springer Nature.

Rittle-Johnson, B. (2019). Iterative development of conceptual and procedural knowledge in mathematics learning and instruction. In J. Dunlosky & K. A. Rawson (Eds.), *The Cambridge handbook of cognition and education* (pp. 124–147). Cambridge University Press.

# IMPACT OF PROMPTS ON EXPECTANCIES FOR SUCCESS, TASK VALUES, AND COSTS IN PROBLEM POSING

Janina Krawitz[1], Katharina Wiehe[2]

[1]University of Paderborn, Germany; [2]University of Münster, Germany

*Recent research has shown that problem-posing prompts affect students' achievement-related outcomes in problem-posing tasks. This study extends such findings by investigating the effects of problem-posing prompts on students' motivational outcomes. Ninth- and tenth-graders (N = 78) were prompted to pose easy and difficult problems. Subsequently, each student reported their expectancy for success, task values, and perceived cost in relation to posing easy versus difficult problems. The results revealed that posing easy compared with difficult problems positively affected expectancy for success, utility value, attainment value, and perceived cost but not intrinsic value. An implication of this study is that including the prompt to pose easy problems in problem-posing tasks is important for students' motivation.*

Problem posing is a promising teaching method for enhancing mathematics learning. In addition, building problem-posing competence is considered to be an important goal of mathematics education (Liljedahl & Cai, 2021). The potential and importance of problem posing were acknowledged in a recent debate, as manifested in an increased number of publications and special issues that have addressed this topic. However, the number of intervention studies has remained notably low, particularly in research on secondary school students (Lee, 2020). An important question in problem-posing research is how different problem-posing prompts affect the process and outcomes of problem posing (Cai et al., 2023). Recent research has indicated that prompting students to pose problems with varying levels of difficulty (e.g., easy, medium, and difficult) yields better outcomes in terms of higher complexity in self-generated problems than prompting them only to pose three problems (Cai et al., 2023). Prompting students to pose one difficult problem might be enough to trigger better outcomes, but the prompt to pose easy problems might be important for their motivation. Hence, the aim of the current study was to address the research gap on how prompting easy versus difficult problems affects students' motivation, including their expectancies for success, task values, and costs in problem posing.

## PROBLEM POSING, EXPECTANCIES FOR SUCCESS, TASK VALUES, AND COSTS

### Problem posing and problem-posing prompts

According to Silver (1994), problem posing involves generating new problems and reformulating given problems before, during, or after problem solving. Our focus here is on problem generation before problem solving. A great variety of different types of problem-posing tasks is have been used in research. Generally, a problem-posing task comprises two different elements: a problem situation and a problem-posing prompt



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 161-168). PME.

(Cai & Hwang, 2023). The problem situation offers a context and data for posing problems, whereas the problem-posing prompt specifies how to work with the problem situation. Problem situations can be either intra-mathematical (e.g., geometrical figures, sequences, equations) or real-world situations (e.g., descriptions of real-world situations, photographs, artefacts). Similarly, problem-posing prompts can pertain to posing either intra-mathematical or real-world problems. Here, we take the perspective of modelling and applications in mathematics education (Niss & Blum, 2020) and focus on problem posing where both the problem-posing situation and the prompt are related to the real world. We refer to this type of problem posing as modelling-related problem posing (Hartmann et al., 2023). Figure 1 presents a modelling-related problem-posing task.

**Freshtival**

Freshtival is a music festival in Holland that attracts music fans from all over the world. To immerse themselves in the colorful world of Freshtival for an entire weekend, music fans can camp on the 120-hectare festival grounds. The Freshtival took place for the first time in May 2006 with 3,000 music fans. Since then, the number has risen steadily to 20,000 music fans in 2017 and even 60,000 in 2022.

The prices for Freshtival 2023 are shown in the table below.

| | Early Bird | Standard |
|---|---|---|
| 3-Day ticket (Friday to Sunday) | 115,50 € | 132,50€ |
| Day ticket Saturday | - | 60,50 € |
| Day ticket Sunday | - | 63,50 € |

Here, you can see the information about Freshtival. Pose an easy and a difficult mathematical problem for your classmates.

Figure 1: Modelling-related problem-posing task “Freshtival.”

In addition to their connection to the real world, prompts can differ in terms of their openness, including open prompts (e.g., “Pose a problem”) and more specific prompts, such as the prompt to pose problems with varying levels of difficulty in the Freshtival task in Figure 1. Recent research has shown that the use of different problem-posing prompts affects learners’ problem-posing achievement-related outcomes, such as the complexity of their self-generated problems (Cai et al., 2023). To the best of our knowledge, the effect of problem-posing prompts on students’ motivational outcomes has not yet been investigated.

### Expectancies for success, task values, and costs in modelling-related problem posing

Expectancy-value theories postulate the central role of students' expectancies for success and the subjective values students attribute to academic tasks in shaping their motivation, performance, and educational choices (Eccles & Wigfield, 2020). Consequently, interventions aiming to enhance students' motivation should address both their expectancies for success and their task values. Expectancies for success are defined as individuals' beliefs about their anticipated performance on upcoming tasks, whereas task values refer to the personal importance of the tasks (Eccles & Wigfield, 2020). Expectancies for success and task values are situated, which means that the expectancy for success and the value of the task vary within the same person across different tasks and different situations. Eccles and Wigfield (2020) distinguished between four types of task values: intrinsic value (interest and enjoyment derived from the task), attainment value (personal/identity-based importance of the task), utility value (importance of the task for present or future goals), and cost (negative aspects of task engagement).

One important factor that affects expectancies and values is a task's complexity. Learners' perceptions of a task's demands impact their expectancies for success and task values (Eccles & Wigfield, 2020). The perception of high difficulty can lead to lower expectancies for success and higher perceived costs. Given the positive association between task values and expectancies for success, a negative relationship between perceived difficulty and task values can be expected (Eccles & Wigfield, 1995). Consistent with these theoretical considerations, previous studies demonstrated that pre-service teachers (Böswald & Schukajlow, 2022) and ninth- and tenth-graders (Krawitz & Schukajlow, 2018) reported lower expectancies for success and lower task values for solving modelling problems compared with less complex intra-mathematical problems, which comprised the same mathematical content but did not require information to be transferred between real-world situations and mathematics.

However, for intrinsic value, the relationship with task complexity might be different. Prior research suggested a positive relationship, as the appraisal of novelty and complexity can evoke interest (Silvia, 2008), and in empirical studies, the complexity of reality-based mathematical problems has been identified as an important factor that contributes to intrinsic value, for both university students (Fielding et al., 2022) and primary-school (Russo & Hopkins, 2017). Students referred to problem complexity as the second most common reason—after the problem context—for inducing their interest (Fielding et al., 2022). Thus, intrinsic value may be enhanced through tasks that that are perceived as challenging.

## PRESENT STUDY AND RESEARCH QUESTIONS

The present study was conducted within the framework of the Mathematical Modelling with Problem Posing (MoPro) project, which is aimed at investigating how problem posing affects affective and cognitive aspects in the context of mathematical modelling. The research questions in the present study were:

How does prompting students to pose easy problems, in comparison with difficult problems, affect (1) their expectancies for success, (2) their task values, and (3) their perceived costs?

To address these research questions, we set up the following hypotheses:

H1: On the basis of expectancy-value theory (Eccles & Wigfield, 2020), we expect that prompting students to pose easy problems will result in higher expectancies for success compared with prompting them to pose difficult problems.

H2: Building on the positive relationships between task values and expectancies for success (Eccles & Wigfield, 1995), we expect that prompting students to pose easy problems will lead to higher utility and attainment values. However, for intrinsic value, we refrain from making a specific hypothesis, acknowledging that posing difficult problems may evoke interest (Fielding et al., 2022; Russo & Hopkins, 2017), which might counteract the expected positive relationships between expectancies for success and task values.

H3: We expect that prompting students to pose easy problems will result in lower perceived costs compared with prompting them to pose difficult problems, as perceptions of task demands are positively associated with perceived costs (Eccles & Wigfield, 1995).

## METHOD

### Participants and procedure

The sample comprised 78 ninth- and tenth-graders from two middle track and one high track school in Germany (50% female; 48.7% male; 15.6 years of age). Each student received a test booklet with four descriptions of real-world situations used as problem-posing situations (see Figure 1). For each situation, students were prompted to generate one easy problem and one difficult problem ("Here, you can see the information about [name of task]. Pose an easy and a difficult mathematical problem for your classmates."). After generating an easy and a difficult problem, students were asked to answer a questionnaire, which was part of the test, about their expectancies for success, task values, and costs related to generating the easy and difficult problems for the given situation.

### Measures

To measure expectancies for success, task values, and costs, we adapted well-evaluated scales from prior studies. Students indicated on a 5-point Likert scale the extents to which they agreed with statements that referred to the constructs (1 = not at all true, 5 = completely true). The following statements were used to measure expectancies for success ("I am sure that in the future I will be able to create a(n) [easy/difficult] problem on the subject of [name of task]"), intrinsic value ("I find it interesting to create a(n) [easy/difficult] problem on the subject of [name of task]"), utility value ("I find it useful to be able to create a(n) [easy/difficult] problem on the subject of [name of task]"), attainment value ("I find it important to be able to create a(n) [easy/difficult] problem on the subject of [name of task]"), and cost ("Creating a(n) [easy/difficult] problem on the subject of [name of task] exhausted me"). For each construct, we aggregated the responses with regard to posing easy problems and again with regard to posing difficult problems across the four problem situations into one scale consisting of four items each, leading to five scales for easy and five scales for difficult problems. The internal consistencies were satisfactory for all scales (Cronbach's $\alpha \geq .757$).

**Treatment fidelity**

In addition, we analyzed the complexity of the tasks to *check the treatment fidelity. Self-generated problems were coded as "easy model" when solving them required a simple mathematical model (e.g., basic mathematical operations only) and as "difficult model" when more advanced mathematical concepts needed to be applied (e.g., Pythagorean theorem or percentages). Non-mathematical problems or missing problems were coded as "no model." For the treatment check, we analyzed the complexity of the self-generated problems that were generated in response to the prompts "easy" and "difficult" and compared the proportions. McNemar-Bowker tests showed that the proportions of self-generated problems for "easy prompts" and "difficult prompts" were significantly different with regard to their complexity for all tasks (all ps < .05) except for Task 2 (p = .572) (see Table 1). A post hoc analysis indicated that significantly more complex problems were generated when the prompt was to pose difficult problems for Tasks 1, 3, and 4 (all ps < .01).* As a consequence, we decided to conduct additional analyses that excluded Task 2, and we report the findings in the Results section.

| | Task 1 | | Task 2 | | Task 3 | | Task 4 | |
|---|---|---|---|---|---|---|---|---|
| | Easy | Dif. | Easy | Dif. | Easy | Dif. | Easy | Dif. |
| No model | 34 | 33 | 17 | 16 | 30 | 36 | 44 | 38 |
| Easy model | 30 | 17 | 3 | 2 | 44 | 21 | 21 | 8 |
| Difficult model | 12 | 27 | 57 | 59 | 4 | 21 | 13 | 32 |

Table 1: Proportions of problems with different complexity generated for the prompts "Easy = Posing an easy problem" and "Dif. = Posing a difficult problem."

***Data analysis***

*Paired-samples t tests were conducted to evaluate whether the expectancies for success, task values, and costs differed for posing easy compared with difficult problems within students. We report one-tailed p-values for the directional hypotheses (H1, H2 – utility and attainment value, H3) and two-tailed p-values for the bi-directional hypothesis (H2 – intrinsic value).*

## *RESULTS*

*The results indicated that, in line with H1, learners reported a tendency toward significantly higher expectancies for success when posing easy compared with difficult problems (*$M_{easy}$ *= 2.96,* $SD_{easy}$ *= 1.19,* $M_{diff}$ *= 2.81,* $SD_{diff}$*= 1.18; t(66) = 1.65, p = .052,* $d_{Cohen}$ *= 0.56). In line with H2, positive effects of prompting students to pose easy problems on utility value and attainment value were found. Students reported that they find it more useful (utility value) (*$M_{easy}$ *= 2.29,* $SD_{easy}$ *= 1.10,* $M_{diff}$ *= 2.08,* $SD_{diff}$ *= 0.93; t(66) = 3.19, p=.001,* $d_{Cohen}$*= 0.53) and more important (attainment value) (*$M_{easy}$ *= 2.10,* $SD_{easy}$ *= 1.12,* $M_{diff}$ *= 1.89,* $SD_{diff}$ *= 0.89; t(66) = 2.46, p = .009,* $d_{Cohen}$*= 0.70) to pose easy problems than difficult problems. With regard to intrinsic value (H2), posing easy problems was not perceived as either more or less interesting than posing difficult problems (*$M_{easy}$ *= 2.16,* $SD_{easy}$ *=1.17,* $M_{diff}$ *= 2.21,* $SD_{diff}$ *= 1.15; t(66) = –0.95, p = .348,* $d_{Cohen}$*= 0.47). For cost, in line with our expectations (H3), learners reported higher costs in posing difficult compared with easy problems (*$M_{easy}$ *= 1.79,* $SD_{easy}$ *= 0.94,* $M_{diff}$ *= 2.02,* $SD_{diff}$ *= 1.03; t(66) = -2.88, p = .003,* $d_{Cohen}$*= 0.64). Figure 2 presents the results.*

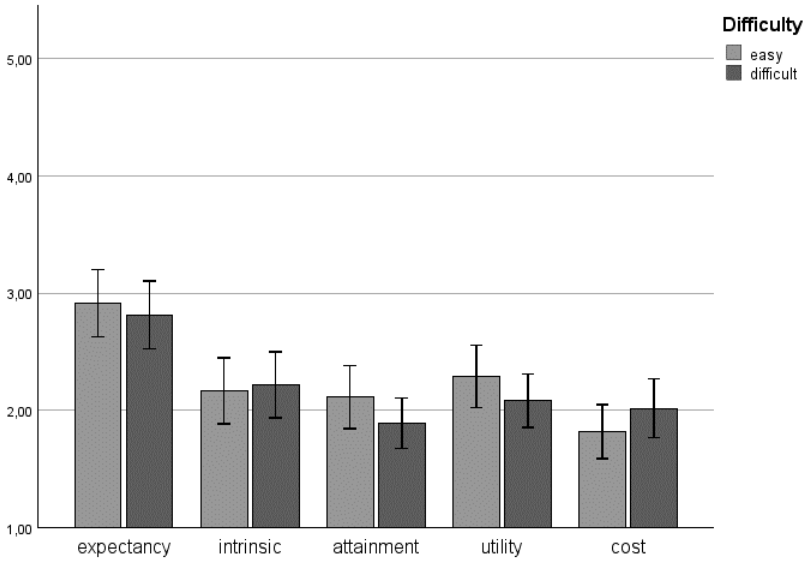



Figure 2: Means for expectancy for success, intrinsic value, attainment value, utility value, and cost for posing easy and difficult problems. Error bars represent 95% confidence intervals for the means.

Because our treatment check showed that, for Task 2, students did not pose significantly more difficult than easy problems, we repeated the analysis without Task 2. The results supported the findings from the previous analysis. The only significant change was that after excluding Task 2, a significant effect of posing difficult compared with easy problems on intrinsic value was found. Posing difficult problems was perceived as more interesting than posing easy problems (*$M_{easy}$ *= 2.12,* $SD_{easy}$ *= 1.20,* $M_{diff}$ *= 2.23,* $SD_{diff}$ *= 1.20; t(66) = –2.11, p = .039,* $d_{Cohen}$*= 0.40).*

## *DISCUSSION*

The aim of our study was to investigate how prompting students to pose problems with varying levels of difficulty affected students' motivational outcomes, specifically their expectancies for success, task values, and perceived costs.

Previous research highlighted the importance of problem-posing prompts on achievement-related outcomes (Cai et al., 2023). Our results extend these findings by demonstrating that problem-posing prompts affect not only achievement but also motivational outcomes. Consistent with our hypotheses, our findings indicate that prompting students to pose easy problems leads to a higher expectancy for success and a lower perception of cost compared with difficult problems. Additionally, students perceived the ability to generate easy problems as more useful and important than the ability to generate difficult problems. These results indicate a negative relationship between the perception of difficulty and the value components attainment value and utility value, supporting theoretical considerations from expectancy-value theory (Eccles & Wigfield, 1995).

For intrinsic value, we found no relationship, or, in the additional analysis, we found a positive relationship when prompting students to pose difficult compared with easy problems. In line with previous research that identified complexity as a source of interest (Fielding et al., 2022; Russo & Hopkins, 2017), students seem to perceive that posing difficult problems is more interesting than posing easy problems. This finding supports the theoretical assumption that the perception of complexity can contribute to interest (Silvia, 2008). Further, the difference between intrinsic value and the other value components highlights the need for a differentiated view of the task value construct and the need to assess various task value components.

A practical contribution from our findings is that prompting students to pose easy problems can help increase their motivation in terms of higher expectancies for success, utility value, and attainment value and lower costs. Solely prompting students to pose difficult problems might be sufficient for positively affecting their problem-posing achievement but is less beneficial for their motivation. This finding is particularly important given the low ratings for task values in our sample, indicating the necessity for improvement.

Acknowledgments: This study was financially supported by the German Research Foundation (GZ: KR 5721/1-1).

### References

Böswald, V., & Schukajlow, S. (2022). I value the problem, but I don't think my students will: preservice teachers' judgments of task value and self-efficacy for modelling, word, and intramathematical problems. *ZDM – Mathematics Education*. Advance online publication. https://doi.org/10.1007/s11858-022-01412-z

Cai, J., & Hwang, S. (2023). Making mathematics challenging through problem posing in the classroom. In R. Leikin, C. Christou, A. Karp, D. Pitta-Pantazi, & R. Zazkis (Eds.), *Mathematical challenges for all*. Springer.

Cai, J., Ran, H., Hwang, S., Ma, Y., Han, J., & Muirhead, F. (2023). Impact of prompts on students' mathematical problem posing. *The Journal of Mathematical Behavior*, 72, 101087. https://doi.org/10.1016/j.jmathb.2023.101087

Eccles, J. S., & Wigfield, A. (1995). In the mind of the actor: The structure of adolescents' achievement task values and expectancy-related beliefs. *Personality and Social Psychology Bulletin*, 21(3), 215–225. https://doi.org/10.1177/0146167295213003

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(4), 101859. https://doi.org/10.1016/j.cedpsych.2020.101859

Fielding, J., Geiger, V., Miller, J., Bruder, R., Towara, U., & Ratnayake, I. (2022). Using situated expectancy value theory to explore initial teacher education students' motivation to engage with challenging mathematical tasks. *Teaching and Teacher Education*, 113, 103663. https://doi.org/10.1016/j.tate.2022.103663

Hartmann, L.; Krawitz, J.; Schukajlow, S. (2023): Posing and Solving Modelling Problems—Extending the Modelling Process from a Problem Posing Perspective. In: *Journal für Mathematik-Didaktik* (44), S. 533–561. https://doi.org/10.1007/s13138-023-00223-3

Krawitz, J., & Schukajlow, S. (2018). Do students value modelling problems, and are they confident they can solve such problems? Value and self-efficacy for modelling, word, and intra-mathematical problems. *ZDM – Mathematics Education*, 50(1), 143–157. https://doi.org/10.1007/s11858-017-0893-1

Lee, S.-Y. (2020). Research Status of Mathematical Problem Posing in Mathematics Education Journals. *International Journal of Science and Mathematics Education*. https://doi.org/10.1007/s10763-020-10128-z

Liljedahl, P., & Cai, J. (2021). Empirical research on problem solving and problem posing: a look at the state of the art. *ZDM – Mathematics Education*, 53(4), 723–735. https://doi.org/10.1007/s11858-021-01291-w

Niss, M., & Blum, W. (2020). *The Learning and Teaching of Mathematical Modelling*. Taylor & Francis Group. https://doi.org/10.4324/9781315189314

Russo, J., & Hopkins, S. (2017). Student reflections on learning with challenging tasks: 'I think the worksheets were just for practice, and the challenges were for maths. *Mathematics Education Research Journal*, 29(3), 283–311. https://doi.org/10.1007/s13394-017-0197-3

Silver, E. A. (1994). On Mathematical Problem Posing. *For the Learning of Mathematics*, 14(1), 19–28. https://www.jstor.org/stable/40248099

Silvia, P. J. (2008). Interest: The Curious Emotion. *Current Directions in Psychological Science*, 17(1), 57–60. http://www.jstor.org/stable/20183249

# EFFECTS OF REFLECTION PHASE TIMING ON PRE-SERVICE MATHMATICS TEACHERS' DIAGNOSTIC PROCESSES

Stephanie Kron[a], Daniel Sommerhoff[b], Kathleen Stürmer[c], Stefan Ufer[a]

[a]LMU Munich, [b]IPN Kiel, [c]University of Tübingen

*Research on pre-service teachers' diagnostic competence revealed that they could benefit from simulation-based learning environments to foster their diagnostic competence. It is emphasized that the diagnostic processes leading to diagnostic judgments should be investigated to understand the development of diagnostic competence. Instructional support, implemented in the simulation-based learning environment, is assumed to affect diagnostic processes positively. This contribution investigates the effects of different timings of reflection phases (concurrent versus final) on the quality of diagnostic processes of N=66 pre-service mathematics teachers. Results reveal that effects of reflection phases on the quality of the diagnostic process differ in accordance with its timing.*

## INTRODUCTION

The key role of diagnostic competence for adaptive teaching and effective learning is increasingly discussed in teacher education research (Herppich et al., 2018). Research on diagnostic competence has turned from solely investigating the accuracy of diagnostic judgements towards also considering teachers' diagnostic processes (Schons et al., 2023). In particular, research has found that pre-service mathematics teachers have difficulties to select suitable, diagnostically sensitive mathematics tasks to elicit student thinking (Kron et al., 2021) and to interpret the observed student solutions. To develop teachers' diagnostic competence, simulation-based learning environments (Heitzmann et al., 2019) as examples of so-called Approximations of Practice (AoPs; Grossman et al., 2009) have been proposed. While these environments turned out to be effective for the development of complex skills in general (Chernikova et al., 2020b), it is still unclear, which instructional support facilitates competence development best in these environments. For simulations in medical education, instructions for reflective reasoning have shown positive effects (Mamede & Schmidt, 2017). One decision to take here is whether reflection should occur at the end of a simulated diagnostic process (final reflection), or if the diagnostic process is paused in-between and resumed after a reflection of the initial process (concurrent reflection). This paper investigates effects of this reflection timing on participants' diagnostic processes when learning to diagnose in a simulated diagnostic one-on-one interviews.

### Teachers' diagnostic competence and diagnostic processes

Heitzmann et al. (2019) define diagnosing as the goal-directed collection and integration of information [for the purpose of reducing uncertainty] to make educational decisions. Diagnostic competence describes the skill to attain high diagnostic accuracy (i.e., agreement between e.g., teachers' evaluation and actual



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 169-176). PME.

student characteristics) over a specified range of diagnostic situations (e.g., having to diagnose student understanding on decimals). Beyond accuracy as the final outcome of engaging with a diagnostic situation, Hammer and Ufer (2023) propose to additionally focus on the processes that lead to the particular diagnostic outcome to assess teachers' competences. Regarding diagnostic competence, this implies investigating the characteristics of the diagnostic processes that are most likely to lead to high diagnostic accuracy. Since a main prerequisite to diagnose students' mathematical understanding is to generate evidence on students' mathematical understanding (Heitzmann et al., 2019), past research has focused on the questions and tasks teachers use to elicit student thinking (Schack et al., 2013). Kron et al. (2021) show that pre-service teachers vary systematically in how well they can identify tasks that potentially contribute to generating diagnostic evidence. Better professional knowledge allows them to identify (and disregard) tasks with low diagnostic potential (i.e., tasks that can be solved correctly even with incorrect strategies and thus do not necessarily provide diagnostic evidence). However, even if a task with low diagnostic potential was selected, novices may at least notice from the student's solution of that task that it does not deliver substantial evidence of student thinking and proceed to the next task, leading to more remaining time for high potential tasks, or for analysing future tasks more deeply before selecting them for the diagnosis. Thus, not only the ratio of selected high potential tasks (among all tasks selected in an interview) may be an indicator for good diagnostic processes, but also a reduced time spent on low potential tasks (in relation to overall interview duration) or increased time on high potential tasks may point to participants' progress in diagnostic competence.

**Fostering diagnostic competence in simulation-based learning environments**

To develop teachers' complex professional skills, such as diagnostic competence, so called AoPs (Grossman et al., 2009) have been proposed. AoPs aim to rebuild a professional real-life situation in an authentic and immersive manner, simultaneously limiting the complexity of the respective real-life situation. As such, AoPs are means to balance authenticity and cognitive demand of learning environments targeting the application of acquired knowledge in professional practices (Codreanu et al., 2020). To improve diagnostic processes using AoPs, instructional support may allow pre-service teachers to activate relevant knowledge, or to identify sub-optimal decisions, improving current as well as future diagnostic processes (Chernikova et al., 2020a).

**Timing of reflection phases during diagnostic simulations**

While expert teachers can (and often have to) rely on well-connected knowledge structures that allow fast, heuristic thinking, novices primarily require explicit reasoning based on factual knowledge to master professional demands (Fink et al., 2021) such as diagnosing student understanding. Enabling learners to look back and reflect on their actions in an AoP may support these explicit reasoning processes (especially for novices) and, for example, help to correct errors or to identify and correct suboptimal actions and decisions that occurred during a diagnostic process

(Mamede & Schmidt, 2017). In their meta-analysis, Chernikova et al. (2020a) report a positive effect of reflection phases on the development of diagnostic competence. For complex professional practices, such as diagnostic one-on-one interviews, reflection phases can have two effects: If the diagnostic process is interrupted for a reflection phase (*concurrent reflection*), this may allow learners to pause and reflect on their behaviour and thus lead to an immediate improvement of the rest of the running diagnostic process. In this sense, concurrent reflection is related to Schön's (1983) idea of reflection-in-action which takes place at "the zone of time, in which action can still make a difference to the situation" (p. 62). It is an open question, if this improved within-training performance would also transfer to subsequent diagnostic processes. Contrary, reflecting after a complete diagnostic process (*final reflection*, or reflection-on-action for Schön, 1983) can primarily improve future diagnostic processes of the learner, does not have the disadvantage of interrupting the ongoing diagnostic process, and allows to reflect on a complete diagnostic process more holistically. Results from medical training indicate that reflection phases during the diagnostic process are more beneficial than reflection phases after the diagnosis (Mamede & Schmidt, 2017). We are not aware of comparable studies for teacher education.

## THE PRESENT STUDY

This contribution investigates the effect of the timing of reflection phases (*concurrent* versus *final*) on the quality of pre-service teachers' diagnostic processes over two consecutive learning simulations:

**RQ1:** To which extent does concurrent reflection improve (a) the ratio of high potential tasks, and the relative time spent on (b) high and (c) low potential tasks during an ongoing diagnostic process?

**RQ2:** To which extent does reflection timing affect subsequent diagnostic processes?

## METHOD

### Role-play simulations of diagnostic one-on-one interviews

We gathered data on the diagnostic processes of $N=66$ pre-service mathematics teachers from two German universities during two consecutive role-play simulations of one-on-one diagnostic interviews on the same day. During the diagnostic interviews, the participants took the role of the teacher and selected tasks from a given task set to elicit the simulated student's answers and draw conclusions about his or her underlying understanding of decimal fractions (Marczynski et al., 2022). Each one-on-one interview lasted up to 30 minutes.

The participants worked on two role-play simulations in the teacher role as part of the learning sessions in a larger intervention study. In each role-play, one fellow pre-service teacher played the role of the student and another one observed the whole process without providing feedback. Only data from the participants in the teacher role will be analysed here. The role-play took place face-to-face, but individual tasks and information were presented by a computer-based system on separate screens for each participant. To allow for repeated measurements, two different student case profile descriptions were designed and provided to the participants in the student role. Assignment of profiles to the two simulations was random. Each triad was randomly assigned to one of two experimental groups: In the *concurrent reflection group* (N=33), the reflection phase started after 12.5 minutes interview time, whereas the participants of the *final reflection group* (N=33) reflected after the end of the whole interview.

To compare between the two experimental groups (concurrent versus final reflection), the diagnostic processes were split into two parts so that the first part was before the reflection and the second part was after the reflection in the concurrent reflection group. In the final reflection group, the two parts were split after 12.5 minutes, the time after which the reflection began in the concurrent reflection group. Four participants in each experimental group were excluded from the analysis, since they selected less than two tasks in one of the four interview parts.

A structured reflection script was developed for the reflection phases based on Mamede et al. (2012). It consisted of seven open questions. Participants had seven minutes to make assumptions about the student's understanding of decimal fractions and to justify those assumptions. Moreover, they were asked to argue which diagnostic information is still needed and how this information could be gathered.

**Instruments**

***Quality of the diagnostic process:*** Three process indicators were investigated to assess its quality. First, we recorded which tasks participants selected for the diagnostic interview. The set of tasks which was provided to the participants consisted of 20 tasks with high diagnostic potential and 25 tasks with low diagnostic potential. Before the first interview, the participants were asked to analyse the tasks' diagnostic potential. The relative interview time (in relation to the whole interview duration for the triad), spent on diagnostic tasks with high (resp. low) diagnostic potential was recorded from log files. Note that beyond the time spent on tasks, participants additionally required time to analyse the task set and select the next task. A higher proportion of high potential tasks, more time spent on high potential tasks, and less time spent on low potential tasks are considered to characterise better diagnostic processes.

***Statistical analyses:*** Descriptive methods as well as linear mixed models were used to analyse the data due to its nested structure. The timing of reflection phases (concurrent versus final), the process part (before and after 12.5 minutes), and the number of the learning simulation (first or second), and their interactions were included as fixed factor in all analyses. The participant and the diagnosed student case profile were integrated as random effects, if they contributed to variance explanation.

**RESULTS**

Over all simulations and process parts, on average 55.0% of the selected tasks were high potential tasks, and participants spent 47.9% of their time on high potential tasks and 34.3% of their time on low potential tasks.

**Effects of concurrent reflection on the ongoing diagnostic process (RQ1)**

The ratio of selected high potential tasks increased significantly from the first to the second simulation part ($F(1,167) = 70.3$, $p < .001$). This increase did not significantly differ by reflection timing ($F(1,167) = 0.1$, $p = .79$), indicating no observable immediate effects of concurrent reflection on the ongoing diagnostic process. The same effect pattern was observed for time spent on low as well as high potential tasks.

**Effects of reflection timing on the subsequent diagnostic process (RQ2)**

Over both experimental groups, the ratio of selected high potential tasks ($F(1,167) = 14.6$, $p < .001$) as well as the time spent on high potential tasks ($F(1,167) = 5.8$, $p < .05$) increased significantly, and the time spent on low potential task decreased significantly ($F(1,167)=12.7$, $p < .001$). These changes towards higher quality of diagnostic processes were significantly stronger in the final reflection group than in the concurrent reflection group for time spent on high ($F(1,167) = 4.6$, $p < .05$) and low potential tasks ($F(1,167) = 4.4$, $p < .05$), but not for the ratio of selected high potential tasks ($F(1,167) = 2.9$, $p = .09$).

| *reflection timing* | *ratio high pot. (%)* | | *time high pot. (%)* | | *time low pot. (%)* | |
|---|---|---|---|---|---|---|
| | Sim #1 | Sim #2 | Sim #1 | Sim #2 | Sim #1 | Sim #2 |
| concurrent | 53.43% | 57.0% | 48.2% | 47.9% | 35.4% | 33.4% |
| final | 49.7% | 60.0% | 43.4% | 52.0% | 39.3% | 29.2% |

Table 1: Descriptive data by reflection timing condition and number of simulation.

Independent of reflection timing, the ratio of selected high potential tasks ($F(1,167) = 4.8$, $p < .05$), and the time spent on high potential tasks ($F(1,167) = 4.1$, $p < .05$) increased stronger from the first to the second part of the first simulation, than for the second simulation. This was not significant for time spent on low potential tasks ($F(1,167) = 1.1$, $p = .29$). The three-way interaction between simulation number, part of the simulation, and reflection timing was not significant for any measure ($p$'s > .36).

**DISCUSSION**

The aim of this contribution was to investigate how the timing of reflection phases as instructional support for pre-service teachers during simulated diagnostic interviews affects their diagnostic processes within each simulation and over two consecutive simulations. While within-simulation (first to second part) development was expected to be positively influenced by concurrent reflection, it was unclear which reflection timing design would lead to better development from the first to the second simulation.

Over both experimental groups, our results reveal that participants' diagnostic processes improve within each simulation (first to second part) as well as over two simulations. This increase is stronger for the first learning simulation. However, there is still room for improvement: For example, more than one third of selected tasks still have low diagnostic potential in the second part of the second simulation. Since participants had already participated in a similar (pre-test) simulation before the first learning simulation considered here, this development cannot be solely attributed to novelty effects, indicating that increased practice with simulations goes along with improved diagnostic processes. Even though it remains to be investigated, whether this can be transferred to real diagnostic processes (maybe even on different topics), it underpins the potential of AoPs and, more specifically simulations (Grossman et al., 2009) to support pre-service teachers' competence development.

Regarding the timing of reflection phases, participants with concurrent reflection did not show (statistically significant) improved diagnostic processes in the same simulation after reflecting, compared to participants from the final reflection group. They also did not show better diagnostic processes in the second simulation than their peers engaging in final reflection. This indicates that—other than in medicine (Mamede & Schmidt, 2017)—concurrent reflection did not show the expected effects on participants' diagnostic competence as observable in their diagnostic processes. One reason might be that it is hard for pre-service teachers to draw implications for immediate further action from the reflection phase, even though this was explicitly prompted by reflection questions. This would mean that reflection-in-practice (Schön, 1983) for diagnostic competence might require explicit preparation in pre-service teacher education (cf. Beauchamp, 2015), for example by explicit reflection training. Also, while the first minutes of the interview serve merely to get an overview of the student's performance over a range of different tasks, some participants might had not yet engaged sufficiently in a more goal-directed analysis of student thinking when the concurrent reflection started after 12.5 minutes. Contrary, engaging in final reflection—based on the full diagnostic process—led to a stronger improvement of diagnostic processes towards the second simulation than concurrent reflection. This indicates that reflecting a complete diagnostic process had some added value for the pre-service teachers in our study.

While the descriptive data indicates the same pattern for all indicators of the diagnostic process, this effect of final reflection was not significant for the ratio of selected high potential tasks. One explanation might be that reflection raised participants' awareness of the varying diagnostic potential of tasks but did not (sufficiently) allow them to perceive this potential based on only reading the task (cf. Kron et al., 2021). Instead, these participants might still be reliant on experiencing to which extent a task really unfolds usable evidence of students' thinking.

An obvious limitation of our study is, that we investigated the effects of reflection over a relatively short period of time. This was sufficient to investigate immediate effects of concurrent reflection on the ongoing diagnostic process, but further research on sustained effects of reflection phases would be of interest. Moreover, the thoughts teacher engage in during reflection may vary systematically. Taking into account the content of teachers' reflections would be promising for future research. Here, inter-individual differences should also be related to participants' professional knowledge.

Summarizing, including reflection phases as instructional support during simulation-based learning environments seems to be a promising approach to assist individuals' development of diagnostic competences towards improved diagnostic processes. However, the results indicate that there is not necessarily an immediate path from reflection towards improvement of their ongoing practice. This could mean that reflection phases as studied here can be one part at the start of the curriculum that targets at establishing reflection *in* as well as *on* own actions as a professional practice. Further research is needed to derive specific implications on the potential role, the design, and the implementation of reflection phases in AoPs. To develop pre-service teachers' diagnostic competence, it remains an open question whether other measures of support, such as explicitly providing or activating required professional knowledge on task potential or student thinking may effectively complement reflection phases.

**Acknowledgments**

The research presented in this contribution was funded by a DFG grant and is part of the research unit "COSIMA" (Grant numbers UF59/5-1 and UF59/5-2).

**References**

Beauchamp, C. (2015). Reflection in teacher education: issues emerging from a review of current literature. *Reflective practice, 16(1)*, 123-141.

Chernikova, O., Heitzmann, N., Fink, M. C., Timothy, V., Seidel, T., & Fischer, F. (2020a). Facilitating Diagnostic Competences in Higher Education-a Meta-Analysis in Medical and Teacher Education. *Educational Psychology Review*, *32(1)*, 157-196.

Chernikova, O., Heitzmann, N., Stadler, M., Holzberger, D., Seidel, T., & Fischer, F. (2020b). Simulation-based learning in higher education: A meta-analysis. Review of Educational Research, *90(4),* 499–541.

Codreanu, E., Sommerhoff, D., Huber, S., Ufer, S., & Seidel, T. (2020). Between authenticity and cognitive demand: Finding a balance in designing a video-based simulation in the context of mathematics teacher education. *Teaching and Teacher Education, 95*, 103146.

Fink, M. C., Heitzmann, N., Siebeck, M., Fischer, F., & Fischer, M. R. (2021). Learning to diagnose accurately through virtual patients: do reflection phases have an added benefit? *BMC Medical Education, 21*:523.

Grossman, P., Compton, C., Igra, D., Ronfeldt, M., Shahan, E., & Williamson, P. (2009). Teaching Practice: A Cross-Professional Perspective. *Teachers College Record, 111(9),* 2055-2100.

Hammer, S., & Ufer, S. (2023). Professional competence of mathematics teachers in dealing with tasks in lesson planning. *Teaching and Teacher Education, 132*, 1-13.

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*, 1-24.

Herppich, S., Praetorius, A.-K., Förster, N., Glogger-Frey, I., Karst, K., Leutner, D., Behrmann, L., Böhmer, M., Ufer, S., Klug, J., Hetmanek, A., Ohle-Peters, A., Böhmer, I., Karing, C., Kaiser, J., & Südkamp, A. (2018). Teachers' assessment competence: Integrating knowledge-, process-, and product-oriented approaches into a competence-oriented conceptual model. *Teaching and Teacher Education, 76*, 181-193.

Kron, S., Sommerhoff, D., Achtner, M., Stürmer, K., Wecker, C., Siebeck, M., & Ufer, S. (2022). Cognitive and Motivational Person Characteristics as Predictors of Diagnostic Performance: Combined Effects on Pre-Service Teachers' Diagnostic Task Selection and Accuracy. *Journal für Mathematik-Didaktik, 43*, 135-172.

Kron, S., Sommerhoff, D., Achtner, M., & Ufer, S. (2021). Selecting Mathematical Tasks for Assessing Student's Understanding: Pre-Service Teachers' Sensitivity to and Adaptive Use of Diagnostic Task Potential in Simulated Diagnostic One-To-One Interviews. *Frontiers in Education, 6*:604568.

Mamede, S., & Schmidt, H. G. (2017). Reflection in Medical Diagnosis: A Literature Review. *Health Professions Education, 3*(1), 15-25.

Mamede, S., van Gog, T., Moura, A. S., de Faria, R. M., Peixoto, J. M., Rikers, R. M., & Schmidt, H. G. (2012). Reflection as a strategy to foster medical students' acquisition of diagnostic competence. *Medical Education, 46*(5), 464-472.

Marczynski, B., Kaltefleiter, L. J., Siebeck, M., Wecker, C., Stürmer, K., & Ufer, S. (2022). Diagnosing 6th graders´ understanding of decimal fractions. Fostering mathematics pre-service teachers' diagnostic competences with simulated one-to-one interviews. In F. Fischer & A. Opitz (Eds.), *Learning to diagnose with simulations - examples from teacher education and medical education* (pp. 49-62). Springer.

Schack, E., Fisher, M., Thomas, J., Eisenhardt, S., Tassell, J., & Yoder, M. (2013). Prospective elementary school teachers' professional noticing of children's early numeracy. *Journal of Mathematics Teacher Education, 16*(5), 379-397.

Schön, D. A. (2017). The reflective practitioner: How professionals think in action. Routledge.

Schons, C., Obersteiner, A., Fischer, F., & Reiss, K. (2023). Teachers' Diagnostic Activities during Task-Based Assessments in a Digital Simulation. 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.

# ALGEBRAIC SEEDS FOR GRAPHING FUNCTIONS

Mathías A. López[1], Susanne M. Strachota[1], Bárbara M. Brizuela[1], María del Carmen Pérez-Martos[2], Angela Murphy Gardiner[3], Maria Blanton [3]

[1]Tufts University, [2]Universidad de Granada, [3]TERC

*This case study of one first grade student involves the analysis of three interviews that took place before, during, and after classroom teaching experiments (CTEs). The CTEs were designed to engage children in representing algebraic concepts using graphs. Using a knowledge-in-pieces perspective, our analysis focused on identifying students' natural intuitions and ways of thinking algebraically about a functional relationship represented using graphs. Findings reveal four seeds, two of which were identified in prior studies, and how the activation and coordination of these seeds results in students' production of function graphs.*

## INTRODUCTION

Recent work in early algebra has shown that algebraic representations, such as variable notation (e.g., Blanton et al., 2017; Brizuela, Blanton, Gardiner et al., 2015; Brizuela, Blanton, Sawrey et al., 2015; Dougherty, 2010) and tables (Brizuela et al., 2021), are within the reach of young children and support them in engaging with algebraic reasoning. We study the natural and intuitive ways that young children's engage in interpreting and constructing of function graphs. To focus on students' intuitions we use of the knowledge-in-pieces epistemological framing (e.g., diSessa, 1993). The fundamental assumptions are that learning can leverage natural intuitions and ways of thinking. This framework has been used to research understandings of multiplication (Izsák, 2005, 2022), probability (Wagner, 2006), integrals (Jones, 2013), and early algebra (Levin & Walkoe, 2022).

Levin and Walkoe (2022) introduce the term seeds of algebraic thinking to refer to small chunks of knowledge that become available to students through interaction with their environment that help them make sense of future algebraic experiences. They describe the following features of seeds: formed in early experience, different from school algebra ideas, and neither right nor wrong as they are context-dependent. We believe this framework could provide a perspective on how students' prior experiences come into play when engaging with graphs. Understanding how students' intuitive knowledge influences their understandings of graphs could open opportunities for instruction and curriculum that build on students' prior experiences.

One seed identified by Levin and Walkoe (2022) is the covariation seed, which involves understanding how an increase in one quantity results in an increase in another. This seed helps students make sense of the effects of the dependent and independent variables in a causal relationship, such as a functional relationship (Levin & Walkoe, 2022). Levin and Walkoe (2022) presented real-life experiences that they hypothesized might be associated with the development of this seed, such as observing a bathtub fill with water. As our research focused on graphs, other intuitive knowledge



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 177-184). PME.

seeds, which have been discussed in previous literature, were likely also activated. For example, "what you see is what you get" (Elby, 2000), which captures when an individual interprets a representation or elements of a representation in a literal sense. For instance, we observed how a student interpreted the points in the graph as actual birds instead of a coordinate pair.

Following Levin and Walkoe's (2022) work, we seek to identify the seeds that are activated when graphing a functional relationship and illustrate how students coordinate these seeds to represent function graphs. We address the following research question:

Which seeds are activated when working with a function graph and how do students coordinate these seeds to construct and interpret a function graph?

## METHOD

We conducted CTEs in Kindergarten and Grades 1 and 2 (ages ranged from 5-8) at an elementary school in the Northeastern United States. In Grades 1 and 2, we taught 14 lessons (see Figure 1). In Kindergarten we taught 16 lessons. We also carried out individual interviews with four students in each of the three grades. Lessons were designed by the research team and based on prior work (e.g., Blanton et al., 2015, 2017; Brizuela et al., 2015). They were taught by a teacher-researcher and were about 30-40 minutes long. All lessons and interviews were video recorded and transcribed. Here we report on three interviews from one Grade 1 student, Lucca.

We selected Lucca's interviews for analysis because his work throughout the three interviews allowed us to construct detailed answers to our research question. All three interviews involved the same questions about the relationship between the number of birds and the number of bird wings, which can be represented as the function $y = x + x$. The students were asked to reason about the relationship and interpret tabular and graphical representations of the relationship or to construct these representations themselves.

Lucca's interview videos and transcripts were reviewed by three team members using microgenetic learning analysis (Fazio & Stiegler, 2013). We tracked instances in which students used seeds, or their own initial ways of thinking about the function graph. The team did not pre-identify the kind of thinking to track. Rather, we looked for evidence of the students' algebraic thinking (i.e., what they said or did) that indicated they were beginning to reason (conventionally or unconventionally) about the functional relationship or the representation. The team reviewed the transcripts individually and then together, until no new instances were identified.

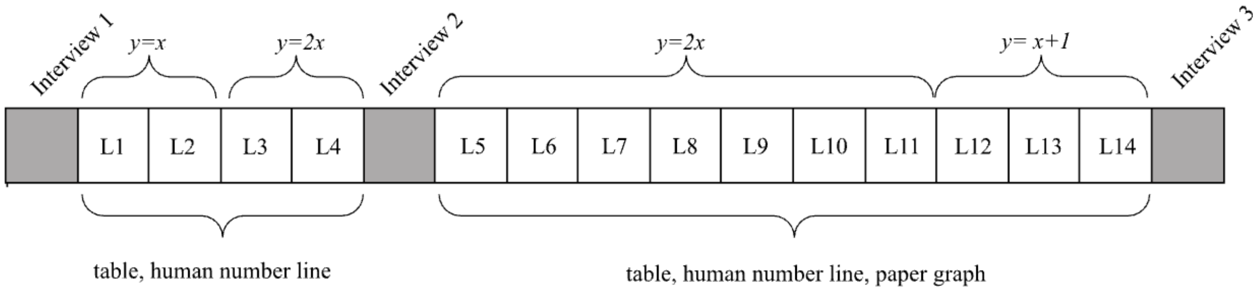



Figure 1. A representation of the Grade 1 and 2 lessons and interviews

## FINDINGS

We observed Lucca use four seeds: *classifying*, *structuring*, *what you see is what you get* (Elby, 2000), and *covariation* (Levin & Walkoe, 2022). Two of those seeds, *classifying*, *structuring,* emerged from our analysis of Lucca's thinking. The other two seeds, *what you see is what you get* (Elby, 2000), and *covariation* (Levin & Walkoe, 2022), were identified previously and then observed in our data. We begin by defining *classifying* and *structuring*.

### Classifying

*Classifying* involves sorting into, identifying, or describing a set. A set is defined by the characteristics of its elements, in this case, all the elements are quantities representing the same variable (i.e., 1 bird, 2 birds, 3 birds, and so on). Lucca likely learned to *classify* early on in real life experiences and through play.

In the interviews, we observed Lucca sort the two variables, "birds" and "bird wings," in the context of a table before activating a covariation seed. That is, before considering how these variables related, he sorted them into two sets by listing the number of birds together and the number of wings together, as depicted in Figure 2.

We also highlight that during the second interview, Lucca determined the number of bird wings for each bird. However, when asked to record the information in a table, Lucca struggled until the interviewer prompted him to add labels, or to classify the sets. It seemed that reasoning about the labels, or naming the sets that he was representing, supported Lucca in identifying and representing the two variables involved. For Lucca, classifying the numbers and naming the sets were precursors to activating a c*ovariation* or a *structuring* seed.

Even though we did not observe an externalization of the *classifying* seed when Lucca worked with the graph, we note that for *structuring* and *covariation* to be activated, *classifying* had to be activated.

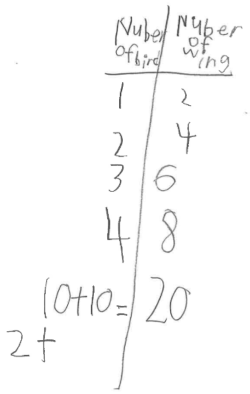



Figure 2. Lucca's self-made table during the third interview

## Structuring

*Structuring* involves coordinating the elements of the sets in a systematic way. Throughout the three interviews, Lucca classified the numbers of birds and numbers of wings and then structured them in a table. In other words, he structured sets when he coordinated the elements in the first column (i.e., the number of birds) and the elements in the second column (i.e., the number of wings), so that he could correctly read across rows.

In the first interview, Lucca was aware of some *structure* linked to the shape of the graph. Specifically, he noticed that the points (1, 2) and (2, 4) were at the intersection of the graph grid lines. However, he did not further *structure* or specifically coordinate the corresponding elements of the sets until later interviews. For example, in the second interview Lucca constructed a self-made graph (Figure 3, left). He connected corresponding quantities with lines but did not plot points until he was prompted to by the interviewer. Once prompted Lucca drew a point on a seemingly arbitrary spot on one of those lines. The following transcript summarizes this conversation and Lucca's representation and the point are shown in the left side of Figure 3.

Interview (I): Where would you put a point to show me two birds have four wings down here?

Lucca (L): It's close to like, almost both of them.

I: Was there a math reason you put it down there?

L: Because I thought it could be anywhere on the line (as seen in Figure 3).

This example shows how Lucca activated *structuring* to coordinate bird and bird wings in a non-canonical way and highlights how seeds are neither right nor wrong since they are context dependent. Lucca's way of *structuring* was likely based on prior experience, he was aware that the point needed to be somewhere on the line, likely because of his prior experiences with the number line during the CTE.

When asked about representing the relationship in his third interview, Lucca correctly labelled the axes. He then explained that those were the correct labels because the *x*-axis corresponds to the "animal" and the *y*-axis corresponds to the "animal (body) part." Furthermore, he was able to plot and interpret the points correctly noting that each point referred to the number of birds and its corresponding number of wings. By the third interview, he could determine the number of birds corresponding to six bird wings by looking at the graph. He did this explicitly by drawing guidelines (see Figure 3, right). Another instance in which we observed Lucca activating the *structuring* seed was when he explained why he knew the location of the points:

I: Why did you put the point right there? What does it mean?

L: Because it should go on the corner.

I: What does it tell me about how many birds and bird wings there are?

L: Because the number of one bird is two bird wings. The number of two birds is four bird wings.

Next, we discuss two seeds that were identified previously in literature and observed in Lucca's interviews, *what you see is what you get* and *covariation*.

**What you see is what you get**

We observed the activation of this seed only during the first interview. When Lucca saw the graph, he first noticed the labels on both axes. He then attended to the numbers and when asked about the points he said, "These are probably the birds;" then after being asked about the first point, he added "This dot is probably the first bird." Here, we note that Lucca is interpreting the points as birds and, thus, is unable to see them as a coordinated pair of the number of birds and bird wings. In other words, the points were not indicating a quantity (i.e., the number of birds or the number of bird wings), but rather the birds themselves.

We attribute the activation of this seed to the fact that Lucca had never seen a graph before, therefore, he did not interpret the points as though they existed in a graph context.

**Covariation**

In the second interview, we observed Lucca create a table. Tables, while not the focus of this analysis, were taught in the CTEs and used throughout the interviews. When asked why he wrote the number of birds first, Lucca's response indicated that he used the same direction change *covariation* seed (Levin & Walkoe, 2022) to reason about the relationship. He explained, "You have to put the birds first to know which (one). So you know that it's the number of the birds." The interviewer probed, "So, you know the number of birds? Why did you put that one first? Why not wings first?" And Lucca further explained, "Because then you would probably get confused." Lucca's explanation suggests he activated a same direction change *covariation* seed and that his understanding of the relation between birds and bird wings at that moment was unidirectional. Lucca's unidirectional understanding surfaced again when he was unable to determine the number of birds given two bird wings. When asked the number of birds if there were two wings, Lucca said, "If there were four wings, there would be two birds."

Interestingly, we observed a shift in the direction of his thinking about the relationship when he interpreted this relationship in a graph context. When Lucca constructed his graph, he actually connected the numbers of bird wings (the *y*-axis) with the numbers of birds (the *x*-axis), which can be seen in his self-made graph (see Figure 3, left). In addition, when given a premade graph and asked to show (i.e., point to) the number of bird wings for three birds he answered, "There's no point. That's not possible" and gestured up the *y*-axis. Based on Lucca's response we assume he had shifted the focus of the directionality of his *covariation* seed, and therefore was unable to reason about three birds. Instead, Lucca seemed to be thinking about three wings and responded that it is "not possible" because he knew no number of birds would have three wings.

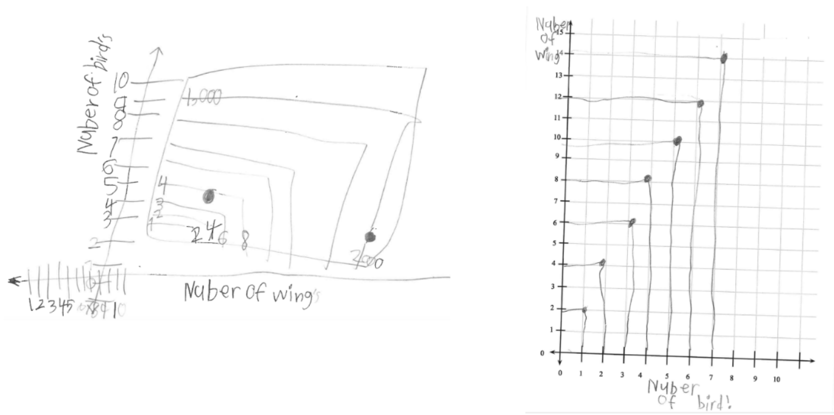


Figure 3. Lucca's self-made graph from his second interview (Left) and Lucca's graph from his third interview (Right)

## DISCUSSION

By Lucca's third interview, we observed him plot a function graph, and we argue that at this point, he was able to do so because he coordinated *classifying*, *structuring* and *covariation seeds*. In other words, we believe that Lucca was able to plot the points because he activated *classifying*, *structuring,* and *covariation seeds* in concert. Even though we did not observe an externalization of the *classifying* seed in the graph context, we note that for *structuring* and *covariation* to be activated, Lucca had to activate and previously engage in *classifying.* Moreoever, we observed several instances in which Lucca engaged in *classifying* in the context of. table, but we do not report those instances here, because there are outside of our focus on graphs.

Note that the *what you see is what you get* seed was likely not activated in these later moments when Lucca graphed because at this point in his development he was no longer relying on that seed. We assume that through his experiences in the CTEs, Lucca became familiar with the function graph, recognizing the relationship being represented rather than just its compelling visual attributes, which are likely to cue the activation of this seed (Elby, 2000).

In the following we briefly summarize our observations of Lucca coordinating the activation of the *classifying*, *structuring*, and *covariation seeds*. We hypothesize that the coordination of the three seeds was a developing ability to activate a *coordination* *class* for representing a functional relationship using a graph. From a knowledge-in-piece framework, a *coordination class* is a task-specific collection of resources students use to engage with the task (Izsák et al., 2022). Here we briefly describe our observations of Lucca beginning to coordinate seeds for graphing functional relationships.

First, Lucca was able to *classify* the number of birds and the number of bird wings in two different sets. He did this by constructing a table listing the number of birds and the number of bird wings together (Figure 2).

He then identified these two different sets in the graph by referring to the labels, and then *structured the sets*, when he described the number of wings as twice the number of birds. The external representation of his *structuring* was evident when Lucca drew guidelines to plot the points, coordinating the increment in $x$ with the increment in $y$. This moment is also evidence that he activated the *covariation* seed because he describes a "resulting change in output" given information about the input (Levin & Walkoe, 2022, p. 1306). Moreover, Lucca plotted points, indicating that he understood how to *structure* the two sets (i.e., he understood how to represent that specific set elements were coordinated).

## CONCLUSION

Different students may activate different seeds in order to graph a function; we do not suggest that Lucca's coordination class will generalize to all students. However, there is significance in analyzing moment-to-moment reasoning and attention to interactions between different seeds to understand students' mental activities. As seen in this work, we identified two elements of Lucca's knowledge which we conceptualized as *classifying* and *structuring*. We believe that these seeds, in coordination with a *covariation* seed supported Lucca in graphing a function.

Future research could focus on exploring how the two seeds we report activate in different problem contexts and representations as well as the possibilities of these seeds refining over time. Additionally, exploring how different students coordinate their seeds to engage in graphing could allow for a better understanding of what experiences incite their activation. Finally, we highlight the potential in using a seeds framework because it supports us in moving "away from the predominant preoccupation with numerical calculations" and placing the "focal emphasis on typical and important ways of mathematical thinking" (Dörfler, 2008, p. 159) many of which are intuitive and natural, based on prior experiences, and captured with the seeds approach to mathematics learning.

### Acknowledgement

This research is supported by the National Science Foundation's DRK-12 Award #1154355. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

## References

Blanton, M., Brizuela, B. M., Gardiner, A. M., Sawrey, K., & Newman-Owens, A. (2017). A progression in first-grade children's thinking about variable and variable notation in functional relationships. *Educational Studies in Mathematics, 95*(2), 181–202. http://www.jstor.org/stable/45184560

Brizuela, B. M., Blanton, M., Gardiner, A. M., Newman-Owens, A., & Sawrey, K. (2015). A first grade student's exploration of variable and variable notation / Una alumna de primer grado explora las variables y su notación. *Estudios de Psicología: Studies in Psychology, 36*(1), 138-165.

Brizuela, B. M., Blanton, M., Sawrey, K., Newman-Owens, A., & Gardiner, A. M. (2015). Children's Use Of Variables and Variable Notation To Represent Their Algebraic Ideas. *Mathematical Thinking and Learning, 17*, 1-30.

diSessa, A. A. (1993). Toward an epistemology of physics. Cognition and instruction, 105-225.

Dougherty, B. (2008). Measure up: A quantitative view of early algebra. In J.J. Kaput, D.W. Carraher, & M.L. Blanton (Eds.), Algebra in the early grades (pp. 389-412). Lawrence Erlbaum Associates.

Dörfler, W. (2008). En route from patterns to algebra: Comments and reflections. ZDM Mathematics Education, 40, 143–160.

Elby, A. (2000). What students' learning of representations tells us about constructivism. *Journal of Mathematical Behavior, 19*, 481–502.

Izsák, A. (2005). "You Have to Count the Squares": Applying Knowledge in Pieces to Learning Rectangular Area. *The Journal of the Learning Sciences, 14*(3), 361–403.

Izsák, A., Beckmann, S., & Stark, J. (2022) Seeking Coherence in the Multiplicative Conceptual Field: A Knowledge-in-Pieces Account. *Cognition and Instruction, 40*(3), 305-350.

Jones, S. R. (2013). Understanding the integral: Students' symbolic forms. *The Journal of Mathematical Behavior, 32*, 122-141.

Fazio, L. K., Siegler, R. S. (2013) Microgenetic Learning Analysis: A Distinction without a Difference: Commentary on Parnafes and diSessa. *Human Development, 56*(1): 52–58.

Levin, M., Walkoe, J. (2022) Seeds of algebraic thinking: a Knowledge in Pieces perspective on the development of algebraic thinking. *ZDM Mathematics Education 54*, 1303–1314.

Wagner, P. A. (2006). Probability, Decision Theory, and a Curricular Approach to Developing Good Thinking. *Journal of Thought, 41*, 23-38.

# EMOTIONAL ASSOCIATIONS WITH MATHEMATICS: USING THE LENSES OF AFFECT AND IDENTITY TO UNDERSTAND PRESERVICE TEACHER STORIES

Ofer Marmur and Lisa Darragh

The University of Auckland

*Key events in one's mathematical learning journey are often recalled with strong emotions and possibly implicated in one's relationship with mathematics. The relationships preservice teachers form with mathematics will, in turn, impact greatly on the way in which they teach the subject in the future; thus an understanding of these relationships is important for the mathematics education field. In this paper we utilise two complementary theoretical lenses, affect and identity, to unpack the written stories of memorable mathematics learning events told by preservice teachers, revealing the deep emotions involved in associating or disassociating with mathematics. We argue the use of each lens enables us a different understanding of the data, yet combined they provide theoretical and practical insight that is greater than the sum of its parts.*

## INTRODUCTION

> "I remember thinking then of mathematics as a kind of omnipotent protector. I was small and quiet and he [the teacher] was large and loud, but I was right and I could show him. [...] Perhaps not surprisingly, the story still evokes the same emotions in me that it did decades ago." (Paulos, 2015, para. 16)

As mathematics educators, there are some common responses we often hear when we meet someone new and tell them what we do as a profession – which include claims such as "oh, I was never a maths person", "it was my most hated subject at school!", or "I really enjoyed maths until the 4th grade". Subsequently, these responses often develop into personal stories people share about what they perceive as key experiences that shaped their mathematical journey, typically associated with strong emotions experienced during those events. Clearly, and has also been demonstrated in prior research, memorable events of mathematics learning as described in autobiographical accounts are implicated in the relationships people form with mathematics (e.g. Martino & Zan, 2010; Towers et al., 2017).

In this paper, we focus on the population of preservice teachers, and examine memorable mathematical events that preservice teachers narrate in relation to their relationship with the subject. Research suggests that preservice teachers at the primary school level often have difficult or anxious relationships with the subject of mathematics (Black et al., 2009), and yet in most cases they will have to teach the subject throughout their careers. It may be useful for those who teach courses in initial teacher education (ITE) to gain further insight into preservice teachers' experiences of mathematical learning to further understand how these may have shaped their current relationships with, and feelings about mathematics, as they prepare to become teachers.



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 185-192). PME.

For our examination, we utilise two complementary theoretical lenses – mathematical affect and mathematical identity. While previous research has either conflated these two domains or else argued for the need to consider them as distinct (e.g. Darragh, 2016), we instead recognise their interrelated nature and see potential value in using both theoretical lenses simultaneously. In doing so, we hope to gain research insight that might not be obtained using either lens separately, with the aim of developing a better understanding of the interconnected roles that mathematical affect and identity play in shaping one's relationship with mathematics.

## THEORETICAL FRAMING

### Affect

Mathematical affect has traditionally been divided into three categories consisting of *emotions*, *attitudes*, and *beliefs*, where this categorisation order represents a decrease in affective involvement and intensity, and an increase in stability (McLeod, 1992). In particular, emotions can be regarded as in-the-moment, rapidly changing, and potentially intense states of feeling that occur during mathematical activity (Goldin, 2000); whereas beliefs are considered to be generally stable individual traits developed over lengthy periods of time, which can be classified as beliefs about mathematics, oneself, mathematics teaching, and social context (McLeod, 1992). Additionally, affect has been shown to be closely interlinked with cognitive processes during mathematical engagement, and it can be manifested in psychological, physiological (embodied), and social ways (Hannula, 2012).

Acknowledging the interconnection between emotion, learning, and memory, Marmur (2019) suggested focusing on students' *memorable events* during mathematics learning – defined as events that hold significance and meaning for the person who experienced them, and are typically accompanied by strong emotions, either positive or negative. Marmur further argued that such emotionally-loaded events that remain memorable in individuals' minds could be used to examine the interrelation between in-the-moment emotions and the formation of longer-term attitudes and beliefs. Here we suggest that such events could also be used to analyse the enactment of one's mathematical learner identity, a notion which we attend to below.

### Identity

A mathematics learner identity can broadly be defined as

> A socially produced way of being, as enacted and recognized in relation to learning mathematics. It involves stories, discourses and actions, decisions, and affiliations that people use to construct who they are in relation to mathematics, but also in interaction with multiple other simultaneously lived identities. This incorporates how they are treated and seen by others, how the local practice is defined and what social discourses are drawn upon regarding mathematics and the self. (Darragh & Radovic, 2018).

For the purpose of this paper, the 'enactment' of identity is the written response to a question about memorable events in learning mathematics. Other 'lived identities' include the preservice teacher identity, which is also enacted through the written response. The 'local practices' of mathematics and of mathematics teaching and learning may be read in the response, and we may also infer how the teacher recognises the learner of mathematics. We draw from the work of Butler (e.g. 1988) to see identity as a "series of acts which are renewed, revised, and consolidated through time" (p. 523), a view of identity that acknowledges its temporal nature. Analysing a written story with an identity lens requires an understanding that the identity is enacted in the moment of writing the story, rather than being a faithful retelling of an earlier identity. The story is of their current relationship with mathematics, which includes their understanding of what mathematics *is* (and how they relate to this mathematics), as well as who the mathematics learner *should be* (and whether this is someone they themselves can be).

### Research question

Utilising the aforementioned interconnected theoretical lenses, we aim to address the following research question: *How are significant, memorable events of mathematics learning implicated in preservice primary teachers' relationships with mathematics?*

## METHOD

For our investigation, we examine written responses of preservice primary teachers (N=59) to a prompt asking them to recall a particularly meaningful memory from their own experience of studying mathematics. In particular, the preservice teachers were prompted to describe why the event was particularly meaningful and memorable for them; address whether the event was influential in the continuation of their mathematics studies; and reflect on whether the event was in some way related to how they thought about their teaching (or future teaching).

As this research project is ongoing, we are still in the middle of the analysis process, which consists of the following stages: (1) iteratively reading the data and identifying themes in the teachers' responses; (2) for each theme, choosing specific data excerpts and analysing them individually – each using our respective theoretical lens (Ofer – affect; Lisa – identity); (3) meeting together to discuss, negotiate, and develop a joint analysis with both the affect and identity lenses; and (4) examining the similarities and differences between what each of the theoretical lenses offers to the data analysis, as well as what potential research insight they can offer combined.

## FINDINGS

Our initial analysis points to four main themes that describe the preservice teachers' memorable events during their mathematics studies: *story of overcoming mathematics*; *experiencing failure*; *the 'horrible teacher'*; and *the 'teacher saviour'*. Due to space constraints, we here present two (shortened) excerpts corresponding to the two latter themes, each followed by two 'readings' drawing on our two theoretical lenses (see Marmur & Darragh, 2024, for additional excerpt examples corresponding to the two former themes). We invite the reader to analyse these excerpts from their own theoretical perspective before reading ours.

**Excerpt 1 (the 'horrible teacher' story)**

> My relationship with math has been a rocky one, recalling a meaningful memory was tough because it brought up a lot of negative experiences that unfortunately influenced my perception and attitude towards the subject. Throughout elementary school I struggled with school, especially math. I would attend […] a tutor once a week, but nothing seemed to help. The stress and embarrassment of not being able to retain and understand the information affected my self-confidence and attitude towards math, I essentially gave up on learning math at a young age. A particular memory that stood out for me was in grade 4 when my teacher had kept me behind during gym because I hadn't finished my division work sheet. I remember her being frustrated at pointing at the problem and trying to explain it to me. I completely shut down, at that point I was too ashamed that I couldn't understand it, I started to feel physical symptoms from the stress, I had a headache and started to feel nauseous. Finally, she became so frustrated that she shooed me off to gym. That evening I went home and told my mum that I didn't want to go back to school and I refused to go for 2 weeks. After my hiatus, I felt withdrawn and physically ill when I entered the class.
>
> Math has always been a source of anxiety for me. However as I am working towards becoming a teacher I have slowly gained confidence in being able to teach it to children. […] As an adult, my perspective on math has shifted and I have gained more confidence through teaching as an [assistant] and being in the teaching program. […] Since I struggled with math so much as a child I think I will be able to relate to my students who are struggling too and be able to support them and give them a positive experience in learning math.

Utilising an **affect** lens to examine the story above, we identify a relationship with mathematics that is extremely emotionally-loaded, riddled with feelings of anxiety and distress. In the introduction to the story, the writer explicitly links repeated negative emotional experiences during their primary school years to the formation of their longer-term attitudes and beliefs towards the subject, as well as their perceived mathematical abilities. In regard to the memorable event, not only is the writer describing psychological manifestations of troubled emotions (e.g. feeling stressed and ashamed), but also physiological and social. That is, the psychological reaction of emotionally 'shutting down' was also accompanied by physical reactions of a headache, nausea, and feeling sick; as well as by a social avoidance reaction of staying home for two weeks.

A particularly interesting point to notice is that the writer does not only describe their own emotions during the experienced event, but also attributes emotions to their teacher. Presumably, without asking the teacher what her emotions were, the writer positively asserts the teacher was feeling extreme frustration, ending with shooing the writer to the gym. While we may not know what the teacher actually felt (and perhaps even the teacher was just kindly trying to help and explain), we can infer that from the writer's perspective, they did not feel emotionally supported at that point in time. This comes in juxtaposition with the later description of the writer's beliefs about themself as a (future) teacher – believing that due to their own struggles with mathematics (such as the ones described in the memorable event), they would be able to support their students and give them a "positive experience in learning math".

Utilising an **identity** lens, we consider this story as describing the writer's current relationship with mathematics and mathematics learning. In the telling of this traumatic story, the preservice teacher appears to be explaining their disassociation with mathematics. The story is about giving up; lines 6-7 state this explicitly, and giving up may also be read in the refusal to go to school for two weeks following the reported incident. Interestingly the teacher is also portrayed as giving up on teaching the writer during the described incident. We might also infer that the preservice teacher is ashamed of their disassociation with mathematics, evident in the description of stress and embarrassment when learning mathematics (line 5) and the shame of the traumatic event itself (line 10).

Regarding the local practices of mathematics and mathematics learning, the story evidences a notion of mathematics teaching as being about a teacher passing on a body of knowledge to a student, who must then retain and understand this knowledge (lines 5 and 10). Such a view of teaching can situate the fault of not understanding with the learner. The common discourse of mathematics as a "struggle" (lines 3, 18, 20) fits well with this view of learning.

Finally, we may also read the writer's teacher identity in this story. There is some work required of the writer in having to align a disassociation with mathematics to their future teaching of mathematics and they manage that conflict by enacting the 'relatable' teacher who can empathise with their struggling students.

**Excerpt 2 (the 'teacher saviour' story)**

> Growing up, my parents were very keen on my development in Mathematics. […] My Dad had always been a natural, his background in engineering was a huge factor in how much emphasis was put on doing well in Math and Science. My Mom on the other hand has shared her horrid stories of being disciplined by teachers for not being as successful. […] Many of my summers in grade school were spent working on my Math skills. Despite all these efforts, I still never felt confident in the subject area. I often would compare my older sister's grades to mine and feel inferior for not understanding numbers the way she did. Although my interest for reading and writing was very evident, I still never felt "smart" because of Math. […]
>
> There was one teacher in particular that really [empathised with] me. Although I was nervous to ask for help in front of my peers, she made her classroom environment a place where students felt encouraged and supported. […] I remember getting really down on myself after a Math test. I got nervous and completely shut down after this. I sat at my desk the entire day twiddling my thumbs, hoping I could hide in my own shell. However, this did not go unnoticed. The next day, I came to class to [an encouraging and inspirational note] left on my desk. [authors: a photo of the note was included in the response]
>
> This was a pivotal moment for me in my Math journey. For the first time ever, I felt seen. I hadn't even realized on my own how much my struggle with Math was getting to me. I spent so long internally convincing myself that I was stupid because I wasn't good at Math,

> that I started dreading school in general. To have this level of encouragement and recognition for trying my best, made me realize that Math was more than just getting the number right. After this, I felt more confident asking for help and taking initiative in my own learning. […]
>
> Till this day, I often go back to this note as a reminder to myself, and the support I hope to give my future students one day. […] Today, my personal attitude towards Mathematics still is insecure and flawed. However, I think this challenge of mine is also imperative in understanding my students. It is something that I have embraced and take as an opportunity to learn, fail and try again. […]

From an **affect** perspective, we can recognise emotions, attitudes, and beliefs similar to those expressed in the first except above. Also here, the writer describes feelings of insecurity and distress, such as being "nervous to ask for help", feeling "inferior" to their older sister's understanding of mathematics, and "dreading school" due to their belief that "I was stupid because I wasn't good at Math". The description even includes physiological manifestations of affective nature similar to those described in the first excerpt – "completely shut[ting] down", while continuously twiddling their thumbs and wishing to hide.

However, the major difference between the two excerpts lies in the writer's feeling of the emotional support that was received from the teacher. The teacher's encouraging note after the test is described here as a "pivotal moment" in the maths journey of the writer – who, "[f]or the first time ever, [] felt seen". This pivotal memorable event both revealed to the writer the extent of the emotions experienced up to that point (a recognition that likely allowed them to begin dealing with this issue), as well as served as a starting point for feeling more confident in their maths learning. Similar to the first writer, also here the writer's beliefs about themself as a (future) teacher is of a teacher who can support their students, and also in this case the writer's own struggles with mathematics are given as the reason they would be able to do so. Though whereas in the first excerpt, this belief was juxtaposed to the memorable event of a teacher who was perceived as unsupportive, here the memorable teacher is used as inspiration for the belief.

From an **identity** perspective, this story draws on a number of societal discourses about mathematics. Firstly that one might be a "natural" (line 2), that confidence can be conflated with competence (line 6), and that being good at mathematics is the same as being smart – irrespective of ability in other subjects (lines 8, 19). It is not hard to imagine that it might be difficult to associate oneself with this natural, confident, and smart type of mathematics learner. However, we see in lines 21-23 a description of mathematics learning that contrasts with that in the first paragraph. Here mathematics is not about getting the answer right, and the confidence required is in asking for help and taking initiative – practices that are available to anyone, not just the "natural", "confident", and "smart" people. Although attributed to the inspirational teacher, we suspect that this preservice teacher had also been influenced by 'local practices' of learning mathematics during initial teacher education that differed from the practices experienced at school, where failing and trying again may be valued.

In this story (and indeed in the previous story also) we can see the role that the recognition by others plays in identity. Rather than being a failure (not natural, not smart), the writer's teacher recognised them as a hard worker who was trying their best (line 21), thus providing an alternative 'script' for being a mathematics learner. This teacher recognition contrasts deeply with the previous story, where the teacher recognised their student as 'not worth it' and gave up on the explanation.

## DISCUSSION

In this paper, we have utilised two distinct yet complementary frameworks – mathematical affect and mathematical identity – to examine how significant, memorable events of mathematics learning are implicated in preservice primary teachers' relationships with the subject. Each lens offers different, yet complementary readings of the stories. From an affect perspective, we can see the strong emotions at play in both the key mathematical learning events and the relationship with the mathematics teacher. From an identity perspective, we can see how the relationship the writer formed with mathematics was tied up in their understanding of what mathematics teaching should involve, and how the mathematics learner should be. Though what we find particularly interesting here are the connections emerging from the two perspectives combined. In the first story, the preservice teacher narrates a disassociation with mathematics and this disassociation is connected to their intense and traumatised emotions. In the second story, the strong emotions are connected to the re-association with mathematics, ostensibly triggered by the teacher's note.

The understandings we may derive from this dual analysis have practical implications in the context of ITE. The affective and identity analysis of the first story indicate two different barriers for this preservice teacher as they head into their teaching career. It is likely they will first need to repair the trauma before they can reassociate with mathematics. Secondly, their pedagogical understanding of mathematics as something that is passed from teacher to student may be problematic considering the writer "gave up on maths" and thus never received the knowledge they must now pass on to their own students. The second story, on the other hand, indicated similar trauma, yet in this case the repair was begun immediately and an adjusted view of what it means to be a mathematics learner is commensurable with their own experience and potential. ITE can thus learn from the second case in order to provide for the first, a finding that resonates with Martino and Zan's (2010) recommendation that learners' emotions and vision of mathematics should be part of the pedagogical knowledge taught in ITE.

From a theoretical perspective, our study aims to offer another step towards a better understanding of the interrelation between affect and identity (see also Heyd-Metzuyanim, 2017). By focusing on memorable events from one's mathematics studies, we could identify interconnections between the emotions experienced (remembered) at the time and the (dis)association with mathematics (i.e. identity). We suggest that the reliability of the account (in the sense of the accuracy of the recalled event, or whether the associated emotions were actually experienced then or intensified afterwards), is of low importance – as ultimately what the prospective teachers recalled and remembered indicates the significance they attributed to the event, which they carry with them into their future and plays a role in their enacted mathematical identity. Accordingly, we argue further research on autobiographical accounts that focus on what people believe to be key moments in their mathematical journey could offer insight into the workings of affect and identity in relation to preservice teachers' relationship with mathematics and help to inform approaches taken during ITE.

## References

Black, L., Mendick, H., Rodd, M., Solomon, Y., & Brown, M. (2009). Pain, pleasure, and power: Selecting and assessing defended subjects. In L. Black, H. Mendick, & Y. Solomon (Eds.), *Mathematical Relationships in Education: Identities and Participation*. Routledge.

Butler, J. (1988). Performative acts and gender constitution: An essay in phenomenology and feminist theory. *Theatre Journal*, *40*(4), 519–531.

Darragh, L. (2016). Identity research in mathematics education. *Educational Studies in Mathematics*, *93*, 19–33.

Darragh, L., & Radovic, D. (2018). Mathematics learner identity. In S. Lerman (Ed.), *Encyclopedia of mathematics education*. Springer International Publishing.

Goldin, G. A. (2000). Affective pathways and representation in mathematical problem solving. *Mathematical Thinking and Learning*, *2*(3), 209–219.

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

Heyd-Metzuyanim, E. (2017). Identity as a nexus of affect and discourse in mathematical learning. *For the Learning of Mathematics*, *37*(3), 33–38.

Marmur, O. (2019). Key memorable events: A lens on affect, learning, and teaching in the mathematics classroom. *Journal of Mathematical Behavior*, *54*.

Marmur, O., & Darragh, L. (2024). Who does it best? Comparing the lenses of affect and identity to explore students' memorable events. *The 15th International Congress on Mathematical Education (ICME-15)*.

Martino, P. D., & Zan, R. (2010). 'Me and maths': towards a definition of attitude grounded on students' narratives. *Journal of Mathematics Teacher Education*, *13*(1), 27–48.

McLeod, D. B. (1992). Research on affect in mathematics education: A reconceptualization. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 575–596). Macmillan.

Paulos, J. A. (2015). *A numerate life: A mathematician explores the vagaries of life, his own and probably yours*. Prometheus Books.

Towers, J., Hall, J., Rapke, T., Martin, L. C., & Andrews, H. (2017). Autobiographical accounts of students' experiences learning mathematics: A review. *Canadian Journal of Science, Mathematics and Technology Education*, *17*(3), 152–164.

# BRIDGING THE GAP BETWEEN RESEARCH AND PRACTICE: EXPLORING A COLLABORATIVE ASSEMBLAGE OF MATHEMATICIANS AND MATHEMATICS EDUCATORS

Yiorgos Mavrommatis and Despina Potari

National and Kapodistrian University of Athens

*In the evolving landscape of undergraduate mathematics education research, the collaboration between mathematicians and mathematics educators has been an area of study. Drawing on Assemblage theory, this research explores the formation of a collaborative group, as process oriented, in the between space of research and practice. The study investigates the affordances and constraints that shape and are shaped by collaborative praxis in the context of a three-year collaboration between mathematics educators and mathematicians. The analysis reveals the complexities, tensions, and potentialities within the collaborative assemblage, offering insights into what such a group can accomplish in the nexus of research and practice.*

## INTRODUCTION

Over the last two decades, within the field of mathematics education research at undergraduate level, the collaboration between mathematicians and mathematics educators has increasingly become a topic of study (Artigue, 2021). Teacher and didactician collaboration, as Robutti et al.'s (2016) review analysis reveals, often concern "aspects of innovation about: mathematical content […], the development of new curriculum […], different pedagogical approaches […], and the integration of new tools and resources" (p. 662). To this direction different forms of collaboration have been established (e.g. Data-extraction agreements, clinical partnerships or co-learning agreements) with different research interest focusing for example on the professional development and learning of participants, action research or common practice development (Kontorovich et al., 2021). Most research related to collaborative efforts between the two communities has focused on the products, while few on the process of collaboration (Bleiler, 2015). Artigue (2021) argues for a thinking not "in terms of dissemination of research results", but rather "in terms of collaborative projects, building and negotiating, jointly with mathematicians [...] that make sense for all those involved, and meet their respective interests and needs". So, a need for investigation of the nature and process of collaboration become focal in order to build on and learn from collaborative efforts between members of the two communities (Bleiler, 2015). To this perspective we pose the questions about how the study of the collaboration process between mathematicians and mathematics educators can inform us about the potential affordances and limitations that research and practice may offer to the development of university mathematics teaching and learning. What realizations and forms of teaching and learning can be developed as research and practice components push and pull the formation of the collaborative group to different "territories".



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 193-200). PME.

## THEORETICAL BACKGROUND

According to DeLanda (2016), the primary ontological uniqueness of Assemblage theory (AT) revolves around the recognition of social entities as legitimate agents. Unlike the exclusive attribution of agency to predefined structures and human entities, AT shifts its focus to conceptualizing social wholes, referred to as assemblages. These assemblages are characterized by their emergent properties and the exercise of capacities by their components. The emergent properties of an assemblage do not stem from the inherent properties of its individual parts but rather from their interactions. It is essential to emphasize the significant role played by non-human components in the formation, functioning, and evolution of these assemblages. An assemblage parametrized through two processes: de/territorialization and de/coding. Territorialization is the process by which the constituent parts generate consensus and alignments that shape a territory of agreement about norms and practices. Deterritorialization is the process by which an established territory destabilized and permits the reorientation of norms and practices that shape a different territory, creating potentialities for new becomings. The second parameter, coding, refers to the rituals, language, and routines of an assemblage. On the other hand, decoding refers to changes in routines, habits, or practices. AT allows to address complex educational phenomena so that “parallel outcomes can be ‘superposed’, making apparently contradictory events equally possible”. (Beighton, C., 2013, p.1296) For example, the complexity of relations between educational roles and the given resources through certain practices can result to contradictory outcomes concerning collaborative efforts. In mathematics education there are several theoretical perspectives that have been adopted to study teacher collaboration with most common the Community of Practice (CoP) (Wenger, 1998). An emphasis in the formation of common goals in the theory of CoP is challenged in AT as the actions and goals can be of multiple directions. Moreover, the duality among ‘internal’ and ‘external’ actors, relations and processes is of no relevance in AT, as most of assemblage’s components constantly cross such boundaries. Thus, within collaborative assemblage tensions are produced (as affordances or/and constraints) which oscillate between research and practice. Under this perspective on one hand agency in the collaborative group becomes a distributed capacity among human and non-human components, that belong not only to humans but also in all other heterogenous material and immaterial elements of the assemblage. On the other hand, as moving away from a human-centre model of research, we need to focus on the complex interactions of elements, not to represent what is done, but more to what these tensions can do. Instead of classifications of practice based on evidence, we call for an elaboration of evidence based on practice formation. This encourages us to answer what else a collaborative group can do in a given milieu by studying practice and avoid trying to fit it into pre-existing views.

## METHODOLOGY

**Data**

This research study took place in a mathematics department of a University in Greece. A group of three mathematics education researchers and five mathematicians was set up on a basis of studying and improving mathematics teaching at the university for three years. Teaching in this department is mainly on a lecture form where the content is presented by the lecturer with no interaction with the students. The studies are too demanding for the students and the majority of students take much longer than the four years of studies to get their degree. Data collection involves several recorded discussions of the group, interviews with students and university staff, university evaluation reports, mathematical texts, and curriculum materials. All the meetings and interviews were conducted in distance mode using Webex or Skype. The data used in this paper to illustrate the analytical process are drawn from our second meeting in 2020, which took place at the beginning of the COVID-19 pandemic when remote teaching became necessary. Following a discussion the first author had with students, he informed the group about their feedback. One of the main themes, was about workshops that could complement lectures to support students' learning.

**Method of analysis**

Based on literature of mathematics education research we can identify three main dimensions upon which collaboration between researchers and practitioners is developing. Thus, resources, forms and participants are the constitutional sets of which relations shape and define collaborative schemes (Borko & Potari, 2020). To this perspective AT provide the conceptual tools to unfold the process of becoming of a such collaborative group. Hence, a focus on process than the products of collaboration can inform us about the conditions under which the gap between research and practice are blurred offering affordances for both. The assemblage of the collaborative group consists of three main classes- the pool of resource units, the plurality of collaborators, the collaborating praxis- of components that, when brought together, shape and define its own being and function. The focus on the sets of relations that operate between the various components of these classes permit us to study the becoming process of the collaborative group as constantly moves in the between space of research and practice. These sets of relations are made up by the resources and collaborators that conjoint through a collaborating praxis which shapes the ongoing form of collaboration. Collaborating praxis refers to the multiple ways the pool of resources and the plurality of collaborators interact through specific relations developed around the goals-activities and thus, offer certain affordances and limitations. The pool of resources refers to cultural and (im)material resources. The plurality of collaborators refers to the educational roles of participants as also the values and ethics with which associate as social actors. These components identified through the data reading and coded as illustrated in the table. The analysis elaborates the collaborative praxis of a group of different professionals of education during the process of assemblage's formation, through the participants' voices. A voice does not represent an individual's identity, but it is a constellation of agentic components, different in nature, that exercise their relational capacities to produce and express meaning. These voices may express different outcomes, they may be supplementary or contradictory and thus push and pull assemblage's formation to different territories, by (re)inscribing different codes of (inter)action.

| Pool of resource units (cultural/(im)material) | The collaborators praxis | The plurality of collaborators |
|---|---|---|
| **Experiences of Students** (*needs and desires concerning learning*) | **Goals-activities** Afforded-Mobilized/ Constrained by: | **Educational Roles**: *Researcher* *Practitioner* |
| **Teaching resources** (*mathematical tasks, mode of students' action and engagement, workshop*) | *Concerns* (about students/ research/ implementation) | *Facilitator* **Values-perceptions of** *Practical mediation of understanding* |
| **Research resources** (*reflective texts, mode of inquiry and reflection*) | *Social/institutional situation challenges* | *(practice-based)* *Theoretical mediation of understanding (research-based)* |
| **Infrastructure** (*space physical-virtual*) | *Teaching/learning considerations* | **Ethics of** *obligation to discipline* *obligation to students* |
| **Digital Tools** (*platforms, digital hands*) | | |

## RESULTS

In what follows we represent each transcript of the participants' voices in chronological order as they are stated during the discussion. For each voice, first we represent the components that constitute it. Second, we outline what kind of initiative, response or reaction the relation of these components frame, based on collaborative praxis by which they are mobilized/afforded or/and constrained. Finally, we offer insights about the possible affordances and limitations that emerge from the interaction of each voice's components. In the end we provide a synthesis of these results.

Yiannis: The students emphasized to me that they think it is very important to do the tutoring workshops because they want to somehow have feedback when they solve an exercise, How can you be sure that what you solved is correct. And that's why they told me that having a tutoring workshop is not for the teachers to solve exercises and students observe but for them to enter the process themselves either individually or in groups to solve and then have someone who can tell them if what they did is correct or what to rethink.

Yiannis's voice consists of the salience of students' needs and desires concerning learning supported by the ethic of obligation to students through a practical mediation of understanding regarding mathematical tasks, mode of students' action and engagement and tutoring workshop. This interaction of the partaking components frames a facilitator's initiative for the collaboration mobilized by concerns and teaching/learning considerations. Specifically, concerns about students' mode of participation/engagement in learning process indicate a counterculture value against conventional lecture and teaching/learning considerations in terms of promoting students' mathematical competence and self-confidence, indicate an ethic of educational success. Thus, a collaborative goal/activity is emerged, with respect to these values/ethics, toward the development of teaching/learning activity through a workshop in relation to students' feedback and support.

> Anna: Students are used to seeing procedures in school and expect the same in university and this confuses them a lot…coming back to something else Yiannis said which I had tried to do but, in the end, didn't even start because of the conditions, the workshop that will give them the possibility of interaction, really solving exercises by themselves and asking about them what they don't have doing well is the most important thing right now, at a distance mode I don't know how it will be.

Anna's voice consists of a relational difference in *students' mode of action and engagement* with *mathematical tasks* based on institutional origin through the ethic of *obligation to* mathematics community *discipline* mediated by the value of *practical understanding*. Moreover, the tutoring workshop interacts with the *mathematical tasks* and *students' mode of action and engagement* supported by the ethic of *obligation to students*. These interactions of the components frame a practitioner's initiative for the collaboration, through goals/actions about the teaching activity in relation to students' feedback and adaption to mathematics discipline that afforded by *teaching/learning considerations* and constrained by *concerns* and social/institutional situation challenges. In particular, teaching/learning considerations offer affordances about how to promote students' mathematical competence and self-confidence through lecture's alternatives, and how to support adaption of students' thinking and action to mathematics community practice form. This is a dual form outcome. On one hand indicate a counter-culture value that seeks lecture's alternative to achieve students' competence as ethic of educational success. On the other hand, it indicates a culture alignment with teaching practice form of mathematics at university in difference with that of school. Meanwhile, these goals/actions seem to be constrained by social situation challenges of pandemic period where teaching needed to become distant, where concerns arise about the implementation of a workshop in this context.

> Liza: Through distant learning, someone can organize a workshop and assign tasks as homework and dedicate some time for a student to discuss how she solved it or what she did when she couldn't. You can't do that for everyone, but you can bring this as an example that may provide feedback to others.

Liza's voice consists of the *virtual space* of a *digital platform* through which *mathematical tasks* and *mode of students' engagement* interact to give form in a *tutoring workshop* supported by the ethic of *obligation to students*, through a *theoretical mediation of understanding*. These interactions of the components frame a facilitator's initiative for the collaboration, through goals/actions about teaching activity in relation to students' engagement and their feedback effectiveness by utilizing digital platforms. These goals/actions are afforded by social challenges, teaching/learning considerations and implementation concerns. Particularly, social challenge of pandemic creates an opportunity for digital platforms utilization, arising productive concerns regarding teaching considerations of workshop's implementation in this context, and thus reflecting an ethic of educational vision. So, teaching/learning considerations are developed around students' active learning and collective knowledge as a counterculture value to department's prevailing forms of instruction.

> Liza: maybe Yiannis and I should make a small text from this small experience with the students and maybe next time we will look at some of these issues from the research side, it would be also good to think of a question that you would like Yiannis to ask the students, something that concerns you, something that you would like to try or have thought about, that you would like to see how the students themselves see it.

In this instance, Liza's voice consists of *student's experiences* along with *reflective texts* that come together under the value of *theoretical mediation of understanding*. Also, other *reflective and inquiry actions* compose this voice, related to practitioners' concerns and knowledge supported by the value of *obligation to* mathematics education *discipline*. These interactions of the components frame a researcher's initiative for the collaboration, through goals/actions about inquiry and reflection upon students' and practitioners' needs and desires related to learning and teaching practices respectively. These goals/actions are afforded by concerns about theoretical understanding of practitioners' pedagogical knowledge and professional learning, in relation to students' needs and desires. This goal/action offers a potential affordance for theoretical knowledge development of teachers' actions, interests, and practices and also, practical knowledge of educational limitations and students' learning difficulties.

> Anna: Alright, I agree, during our discussion this idea came to me, that it could be an opportunity now, perhaps in this situation, to try out these new methods, and do it in small groups, in large ones I am not sure that this could work but if it's something that works it could be used under normal circumstances, necessarily this will work with a small group if everyone suddenly decides to participate we won't have time to do anything else.

Anna's voice, in this instance, consists of *physical and virtual space* of teaching as interrelated with the *tutoring workshop*, supported by the value of *practical understanding*. The interaction of these components frames a practitioner's reaction to previous insights about workshop's implementation and possible limitations. Thus, the collaborative goal/activity of workshop is afforded by teaching/learning considerations regarding the integration of the previous negotiation and suggestions, indicate an ethic of educational vision in terms of experimentation in teaching practice. Also, constrained by concerns about practical challenges of workshop implementation regarding the number of participants, based on institutional challenges stemming from a lack of academic staff, due to insufficient funding.

Liza: you won't discuss all 50 solutions, but you can discuss some of them.

Liza's voice consists of the *physical and virtual space* of instruction as interrelated with the *tutoring workshop*, supported by the ethic of *obligation to students*. The interaction of these components frames a facilitator's response to the workshop's limitations that mentioned previously. Thus, the collaborative goal/activity of workshop is afforded by concerns regarding practicing that expressed as suggestions that seek to overcome previous constraints, indicating an ethic of educational vision.

Anna: what I think about what Liza says if, e.g. you have seen five solutions and you have seen five approaches to them, with the student's permission in a discussion per 10 per 20, as we do now with our digital hands (digital tool for communication), to explain to them what the mistakes are, to give them hints without necessarily having seen all 50 of them

Anna's voice consists of *virtual space*, *digital hands*, *mathematical tasks*, *mode of students' engagement*, and *tutoring workshop* through the value of *practical mediation of understanding*, dependent on the ethic of *obligation to students*. The interaction of these components frames a practitioner's response to Liza's voice regarding a praxis about workshops' organization by utilizing digital tools and the mode of students' engagement for their learning support. This goal/activity is afforded by teaching/learning considerations about the effectiveness of students' learning support through a virtual form workshop by utilizing the offered digital tools like digital hands. Furthermore, is afforded by concerns of implementation in terms of students' mode of engagement and management of the number of participants. Thus, this goal/activity provides affordances for resource development like the virtual workshop to support students' learning by utilizing digital tools, that indicate a value of counterculture and an ethic of educational vision.

So, what the assemblage of collaboration can do?

First, to create common interest toward community building based on agreements and alignments about the development of a supporting structure to students' learning, driven by their needs, the mode of engagement and the social challenges. Second, to create conditions under which both theoretical and practical knowledge will be developed. There is a constant interplay between practical and theoretical understanding within the collaborative dialogue. Participants engage in both theoretical mediations of understanding, reflecting on broader educational concepts, and practical considerations related to the implementation of the tutoring workshop. Third, to create resources of collaboration such as the tutoring workshop with the utilization of digital tools based on affordances and constraints. The affordances for collaboration are seen in the opportunities created by the pandemic to explore new methods, utilize digital platforms, and engage in reflective practices. Constraints include challenges related to the number of participants, limitations in resources, and the need to adapt to the institutional situation, such as insufficient funding. These creations are bound to social ethos and values that participate in and derive from the assemblage's formation.

## DISCUSSION

All these potentialities of/for goals-actions and their related affordances and limitations, are developed through processes of de/territorialization and de/coding and thus pull and push assemblage's formation to different directions. An oscillation between research and practice become apparent as assemblage moves to different territories, inscribing different codes of praxis. Thus, on one hand, the participants are engaged in the process of defining and structuring the tutoring workshop, both in physical and virtual spaces. This includes discussions about organizing workshops, assigning exercises, and utilizing digital platforms for distance learning. These actions represent a form of territorialization, where they are establishing specific territories for the collaborative activity. This territory is characterized by a high degree of homogenization regarding teaching and learning considerations related to practice. On the other hand, there are instances of deterritorialization, where traditional or established ways of teaching and learning are challenged. The shift to small group interactions, the use of digital platforms due to the pandemic, and the consideration of new methods represent deterritorialization as they move away from conventional instructional practices. Deterritorializing aspects loosen the dense of the territory's connections and thus make assemblage more open to research and technology integration by which new forms of teaching and learning can be developed. Thus, the process of de/coding (re)arranges and (re)organizes information, meanings, actions and values that derived from research and practice. While the process of de/territorialization determines the degree of alignment to conventional practice norms and values or the potentialities that research in mathematics education and technology provide concerning the teaching and learning development within collaborative group.

## References


Artigue, M. (2021). Mathematics education research at university level: Achievements and challenges. In V. Durand-Guerrier, R. Hochmuth, E. Nardi, & C. Winsløw (Eds.), *Research and development in University Mathematics Education* (pp. 3–21).

Bleiler, S. K. (2015). Increasing awareness of practice through interaction across communities: The lived experiences of a mathematician and mathematics teacher educator. *Journal of Mathematics Teacher Education*, 18(3), 231-252.

Borko, H. & Potari, D. (Eds.). (2020). Teachers of mathematics working and learning in collaborative groups, *Proceedings of the 25th ICMI Study Teachers of Mathematics working and learning in Collaborative Groups.* Lisbon: University of Lisbon.

DeLanda, M. (2016). *Assemblage theory*. Edinburgh University Press.

Kontorovich, I., & Bartlett, P. (2021). Implementation of research on scriptwriting in an undergraduate mathematics course: A study of teacher-researcher collaboration. *ZDM–Mathematics Education*, *53*(5), 1109-1120.

Robutti, O., Cusi, A., Clark-Wilson, A., Jaworski, B., Chapman, O., Esteley, C., Goos, M., Isoda, M., & Jourbert, M. (2016). ICME international survey on teachers working and learning through collaboration: June 2016. *ZDM Mathematics Education, 48*, 651–690.

# MATHEMATICS TEACHERS' REIFIED IDENTIFYING

Mayumi Kawamura

Oita University

*The aim of this paper is twofold: first, it proposes how the analysis of students' implicit and indirect identifying as found in previous studies, can be used as for the analysis of teachers' identities; second, it seeks the characteristics of actions according to which teachers' identities are manifested. To achieve this aim, I analyzed a trigonometric ratios class that took place in the first year of a Japanese public high school. It was found that indirect verbal identifying was found in the activities set by the teacher for the students, and implicit nonverbal identifying was found in the teaching arts used by the teacher. I also demonstrate that the teacher intentionally set learning opportunities for exploration routine while simultaneously having the students perform a ritual routine.*

## BACKGROUND AND RESEARCH QUESTIONS

Mathematics learning is conducted by people. To capture it, therefore, it is necessary to consider not only cognitive mathematization but also subjectifying (also called identifying), which "occurs when the discursive focus shifts from action and their objects to the performers of the actions" (Sfard, 2008, p. 113). Identifying and identity are used nearly interchangeably in this paper. Focusing on identity, we reconsider the relationship between mathematical learning and its sociocultural context (Sfard & Prusak, 2005a). Here, identity is defined as "a narrative about individuals that reifying, endorsable, and significant" (Sfard & Prusak, 2005b, p. 16). It appears not only in direct utterances but also in implicit and indirect identification (Heyd-Metzuyanim & Sfard, 2012). Heyd-Metzuyanim and Sfard (2012) proposed a framework for analyzing implicit and indirect identifying with students. Because teaching is conducted by the students and the teacher together, it is necessary to examine the implicit and indirect identifying of both the students and the teacher. The research questions for this paper are as follows. Can the perspective of Heyd-Metzuyanim and Sfard (2012) be used to analyze teacher identifying? What are the characteristics of teaching by identifying? Through this investigation achieving this purpose, we can interpret the influence of the teacher's identity on mathematics learning.

## PREVIOUS STUDY

Commognitive theory (Sfard, 2008) considers thinking to be a form of communication. By this means, the theory of commognition overcomes the limitations of traditional cognitive research and provides an objective research method for analyzing all, not necessarily only verbal, communicative action. Furthermore, it enables us to view the teaching and learning process of mathematics from cognitive, affective, and social aspects (Sfard, 2012).



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 201-208). PME.

In recent years, there has been a growing interest in commognitive theory in routines of ritual and of exploration. Ritual routines are process-oriented performances, where "rituals are routines performed for the sake of social rewards or in an attempt to avoid a punishment" (Lavie et al., 2019, p. 166). Examples of ritual routines include connecting with others and following known procedures. Exploration routines, on the other hand, are outcome-oriented; they are conducted "to yield a new 'historical fact', a new 'truth' about mathematical objects, etc." (ibid.). Examples of exploration routines include adaptation procedures.

Identity is also important in commognitive theory of. This is because, to interpret mathematical learning and teaching, it is necessary to focus on both mathematization, which targets numbers and figures, and subjectifying, targeting the performers of actions. Identity signifies subjectifying (Sfard, 2008). In commognitive theory, subjectifying is also called identifying (Sfard, 2008). On this basis, Heyd-Metzuyanim and Sfard (2012) present a classification of direct subjectifying utterances and types of identifying narratives according to the means of reifying, that is, making abstract concepts concrete.

Below, I describe and clarify the classification of utterances and types of narratives proposed by Heyd-Metzuyanim and Sfard (2012). Utterances have there are three levels of classification. The first concerns reference to specific actions, such as "I forgot," or "You said." The second relates descriptions of routine performance, such as "I can't," and "It boggles my mind." Finally, the third level concerns about the actor, such as "My brain is so slow," and "She has a mathematical mind." Heyd-Metzuyanim (2013) states that this third level is "identifying by definition" (p. 345).

Types of narratives (also referred to as "identifying" narratives) are first classified as "verbal" and "implicit nonverbal." Then, "verbal" narratives are classified as "direct" and "indirect" (Heyd-Metzuyanim & Sfard, 2012). Only "implicit nonverbal" and "indirect verbal" identifying are discussed here. Examples of implicit nonverbal narratives are nonverbal symbolic tools, including sounds, gestures, facial expressions, eye contact, and "repetitive, consistent nonverbal actions" (Heyd-Metzuyanim & Sfard, 2012, p. 131). Many of these are linked to verbal communication, but they can also be elements of repeated and consistent nonverbal identifying, such as when "a student regularly groans at the sight of a fraction" (ibid.). There are two types of indirect verbal. First, there is repetitive, consistent, first- and second-level identifying. For example, statements such as "I don't know" or "I don't understand" belong to this type. The other type involves the positioning of the person. This entails a statement that has the form of referring to another person but that is actually about the speaker and indirectly indicates the speaker's position toward the other person. For example, "Why does x do this?" (ibid.).

This study explores teacher identities through the application of the Heyd-Metzuyanim and Sfard framework. I propose to investigate whether their framework is compatible not only with student identity research but also with teacher identity. This study addresses the empirical gap in teacher identity research by incorporating a less commonly used framework.

## METHOD

### Data Collection and Date Analysis

The data in this paper are a video recording and accompanying transcription from the perspective of the back of the classroom. The language of the data was Japanese. The first class was a 50-minute first-year math class at a Japanese public high school, with 40 students. This was a revision class conducted after the students had learned trigonometric ratios. The teacher had a master's degree in education and 11 years of teaching experience.

This study analyzed the transcribed data using qualitative methods. To determine the implicit and indirect identifying of the teacher, the author used the implicit and indirect identifying framework that was proposed by Heyd-Metzuyanim and Sfard (2012). In this analysis, I first identified situations in which correct answers or statements were obtained as intended by the teacher and those where the teacher did not receive an intended response. Next, the characteristics of each situation were classified according to whether they were implicit or indirect. The classification procedure was as follows.

- The basic difference between implicit nonverbal and indirect verbal identifying was determined where the former used verbal expressions and the latter did not.
- I added an additional type of implicit nonverbal identifying based on its intended meaning. Because it is rare to observe class periods completely without language. Here, I also identified implicit nonverbal actions that were part of the intended meanings from their expressions.
- The difference between the two types of indirect verbal identifying was determined by whether the given person was positioned or not (positioned indirect verbal vs. nonpositioned indirect verbal).

From this analysis, I examined the actions of the students resulting from the actions of the teacher to characterize their reactions due to the implicit and indirect identifying of the teacher. In doing so, I used the perspectives of ritual and exploration routines (Sfard, 2008). All of these analyses were conducted in Japanese and were translated into English by the author for this paper.

## ANALYSIS RESULTS

### Sharing Plans for Solving a Problem Reified as Indirect Verbal Identifying

The class began with the following statement by the teacher with respect to the blackboard problem "Find the maximum and minimum values of $y = \sin^2\theta - \cos\theta$ when $0° \leq \theta \leq 180°$."

1 T: What do you think of $y = \sin^2\theta - \cos\theta$? I think that you guys have solved that problem at home. So please share this with your classmates. (Omitted.) If

you don't know what to think at first, some of you may not be making much progress, we'll just plan. Just the plan.

The teacher repeatedly set out opportunities to share plans for solving the problem, as in 1T. In response to these instructions, the students shared plans for the solution of the problem. To check obtained solution as a class, the teacher could have proceeded without sharing the plan with other students. However, the teacher repeatedly set up situations in which plans were shared. This was a habit in this class, after having been done many times. This can be interpreted as a sign of the identity in which the teacher emphasizes the importance of reflecting on problems after sharing the plan. Although in several situations, when the teacher set the students up to share plans for solving a problem with their classmates, he did not explicitly say, "It is important to consider a plan for the solution." This was a representative example of indirect verbal identifying by the teacher.

**Comparison and Valuing of a Problem Reified as Indirect Verbal Identifying**

After giving the students the opportunity to share plans with their classmates, the teacher asked them to compare $y = \sin^2\theta - \cos\theta$ with the quadratic function. Following this, the teacher provided his understanding of the value of the problem and moved on to the next problem, $2\sin\theta < 1$. For the problem situation of inequalities involving trigonometric functions, the teacher set up an activity to discuss differences between the properties of equations and those of inequalities. The students were then asked to check their solutions with other students, and one student was nominated to present his solution.

2 T: First , how do we solve the maximum and minimum? By the way, if this is the case (writing $y = x^2 - x$ on the blackboard), how would you think of the maximum and minimum? (Omitted: nominating students, interacting with them, and showing them the solution.)

3 T: This is a collaboration between quadratic functions and trigonometric ratios that we are thinking about, but when solving these, well, you have to be careful, or rather, you have to be concerned about this first move (pointing at $1-\cos^2\theta$ on a blackboard), don't you? That's right. Well, the point is, in our classes, we have talked a lot about how it is better to reduce variables as much as possible. If there are two variables, it would be better to reduce them to one variable. (Omitted.)

4 T: (When $2\sin\theta < 1$) We did first-semester work on linear inequalities and quadratic inequalities, but do you remember the properties that we used to solve first inequalities, first-order inequalities, and first-order equations? What was different when we solved first-order inequalities from first-order equations? What was different when we solved first-order inequalities from first-order equations? Please share this with your classmates. (Nominates S1.)

5 S1: Be careful with 0.

6 T: Be careful (speaking slowly) with 0? What?

7 S1: Oh, be careful with the minus sign.

8 T: Careful with the minus sign? How? (Listens carefully.) (Omitted: after discussing the differences between the properties of linear equations and linear inequalities and having the students present their ideas, the teacher asks the students to present their ideas on 2sinθ < 1, while he writes the solution on the blackboard.)

9 T: This idea is very important, and it means a lot to think about the standard, sinθ, which is smaller than 1/2. We all discussed drawing a unit circle in the equation.

As seen in 2T and 4T, the teacher repeatedly set up comparisons in class. In such cases, the teacher did not call for a direct comparison between A with B. However, he set up situations in which students could compare quadratic and trigonometric functions and compared the properties of linear equations and inequalities. Although the content of the comparisons differs, the repeated comparison situations can be seen as part of the teacher's emphasis on comparisons. Because the students were already familiar with quadratic functions and inequalities involving trigonometric functions with positive coefficients, the inclusion of comparisons did not assist them in solving the problem. However, the teacher nevertheless insisted on setting up comparisons. This can be interpreted as a sign of the teacher's identity as one who assigns importance to comparisons. Here, the teacher did not provide direct verbal instructions, such as "Let's make a comparison," but he presented the situation of comparison to students, which can be regarded as a nonverbal identifying.

As seen in 3T and 9T, after the teacher nominated students and asked them for their solutions, the teacher did not consider their statements but rather the problems themselves, for example by saying, "It would be better to reduce it to one variable," or "This idea is very important." In addition, generally, teachers can move on to the next problem after solving one problem; however, this teacher did not do that. For each, the teacher conveyed to the students the ideas that he wanted them to acquire, which were not limited only to the problem itself. In this way, each problem was assigned a value by the teacher. This can be interpreted as a sign of the identifying of the teacher, who wants the students to learn how to think in general more than to learn how to solve a particular problem. The teacher did not directly state, "The value of this problem is...," which amounts to a nonverbal identifying of the teacher.

**Writing on the Blackboard, Speaking Rate, and Repetition as Implicit Nonverbal and the Meaning of the Teacher's Utterances as Verbal**

After discussing the inequality $2\sin\theta < 1$, $2\cos\theta+1 \geq 0$ and $\tan\theta-1 \leq 0$ were considered. For the class work on $\tan\theta-1 \leq 0$, the teacher called on student S2, who initially thought that the range of θ was $0° \leq \theta \leq 45°$ but then realized that the slope could be negative, so student (S2) added $90° < \theta \leq 180°$. The following conversation occurred.

10 T: Okay, let's go, S2, go ahead. Tell us about your initial idea.

11 S2: If the slope of tanθ is 1, then θ is 45°.

12 T: (The teacher listens and writes on the blackboard the radius of motion of 45° on the unit circumference.) When the slope represents 1, θ is 45°.

| | | |
|---|---|---|
| 13 | S2: | When it is less than 1, it is between 0° and 45°. |
| 14 | T: | (The teacher, listening, writes.) 0°–45°. (Marked on the unit circumference, $0° \leq \theta \leq 45°$ on the blackboard.) (Omitted) |
| 15 | T: | Earlier, I told you to check the diagram, but where does tanθ appear? sinθ and cosθ are the y- and x-coordinates on the unit circle, right? How about tanθ? S3? |
| 16 | S3: | A slope. |
| 17 | T: | A slope? Anything else? (The teacher slows down and listens.) |
| 18 | S3: | (No reaction) |
| 19 | T: | (Omitted). Where does the slope appear when you draw a graph other than the slope? The unit circle. Ah, tanθ on the unit circle, not on the unit circle, the value of tanθ. Where does the value of the slope appear? Please discuss this, everyone. |
| 20 | S4: | (In group discussion.) Isn't he referring to this line? (Pointing to the line x = 1 drawn in the notebook.) No? |
| 21 | S5: | (In group discussion) I'm not sure what it is. (Omitted.) |
| 22 | T: | Tanθ is a slope, but where does the slope come from? (Omitted. The teacher nominates S6.) |
| 23 | S6: | x = 1. |
| 24 | T: | Yeah. x = 1 and? of the intersection, where that angle is extended from the center? Of the intersection? |
| 25 | S6: | The y-coordinate. |

First, I analyze implicit nonverbal identifying. In 12T and 14T, the teacher listened to the students while discussing and wrote the necessary content on the blackboard. In 10T, the teacher asked the student to "Tell us your initial idea," indicating that the teacher was already of S2's idea thanks to desk-to-desk instruction. The teacher therefore intended in advance to write the student's statement on the blackboard. By contrast, in 17T and 24T, the teacher did not write the students' proposals on the board but repeated the students' statements slowly or, as in 19T, set up an opportunity to discuss again by rephrasing the question. This action shows the students that what they said was not what the teacher intended.

Next, I present indirect identification. In 15T and 19T, the teacher repeatedly asked the students, "Where does tanθ appear?" I interpret the statement made by the teacher to mean "I (the teacher) am not satisfied with the answer, slope, but want you to state that tanθ appears in the figure as the y-coordinate when x = 1." In 19T, the teacher said, "Please discuss, everyone." Here, we find three intended meanings. First, how tanθ appears on the diagram is important, and the teacher asks the students to express this in their own words. The second refers to the teacher's hope that the students would use a different expression than "slope." The third is the teacher's desire to elicit a response by the students by reading the teacher's intention of asking for expressions that are not slopes. I will provide the reasons for each of these below. The first intention finds it reason in that the teacher did not mention x = 1 at this point, while the students explored. Second, the teacher set up the discussion as an opportunity to consider a paraphrase for the term slope. The reason for the third meaning is that the teacher asked S3 at 15T, “Where does tanθ appear?” and the teacher asked the students to provide a non-slope answer at 19T. These implicit actions could be interpreted as hints to the appropriateness of their answers, as well as clues for the intended teaching content that the teacher had prepared for the class.

## DISCUSSION

### Validity of Using the Implicit and Indirect Identifying Perspective in the Analysis of Teacher Identity

As indicated in the previous section, the teacher’s identifying reified itself in both indirect verbal identifying and implicit nonverbal identifying. The former included the comparison and evaluation of problems with the meaning of the teacher’s utterances, and the latter included the action of writing on the blackboard, alterations in the rate of speaking, and the repetition of utterances. All of these were observed several times over the course of a one-hour class. Because the class subject in this study was mathematics, the roles of the teacher and students were clear, and no indirect identification of positions was observed.

I provide a comparison between students and teachers with respect to the types of identifying according to Heyd-Metzuyanim and Sfard (2012). For the students, the implicit nonverbal response was the act of groaning, and the indirect verbal response was repeated utterances such as “I don’t understand.” For the teacher, setting up opportunities for students to review and compare their plans for solving problems, evaluating them, and providing teacher utterances containing different meanings are examples of indirect verbal communication. The writing on the blackboard, the speed with which the teacher asked questions, and the repetition of statements are examples of implicit nonverbal communication and form part of the teacher’s instructional techniques.

### Setting up Ritual and Exploration by the Teacher at the Same Time

In this section, I argue that the teacher set up a ritual routine allowing students to engage in exploration. Such exploration leads to communion with both the mathematical content and with the teacher, deepening students’ understanding. In this study, the teacher provided opportunities for students to ask questions and engage in discussion with regard to the visual interpretation of tanθ. I analyzed the teacher’s utterance at 19T, “Please discuss, everyone” as showing three levels of intention. The first is that the visual interpretation of tanθ is important, and the second refers to the teacher’s recommendation to perform the exploration routine. Consequently, during the discussion opportunity that was set up by the teacher at 19T, S4 asked, “Isn’t he referring to this line?” (20S4), showing the line $x = 1$ that she illustrated in her notebook. Third, by setting up an activity in which he repeatedly discussed the problem, the teacher enabled the students to communicate with the teacher, understanding that he was looking for an answer aside from “slope.” I interpret this as a ritual routine because of the social reward that is obtained by being recognized by the teacher for speaking up. The teacher's may have had several intentions in setting up this activity: perhaps the students were not responding well, or he wanted them to say something important in their own words. In any case, the visual interpretation of tanθ can be understood as an activity in which new truths about mathematical objects are obtained and in which the teacher simultaneously provides an opportunity for the students to explore the teacher's intentions. This is distinct from what Nachlieli and Tabach (2019) proposed, in which a class is held between ritual and exploration. It is also different from the perspective of Christiansen et al. (2023), that a class becomes a ritual, even if the teacher initially sets an exploration, because of the need to conclude a class content within a certain amount of time and the need to balance the consideration for the entire whole student body and the individual students. In this study, the teacher set up ritual and explorative routines to take place simultaneously.

## ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI Grant Number 23K18944.

## References

Christiansen, I. M., Corriveau, C., & Pettersson, K. (2023). Hybrids between rituals and explorative routines: opportunities to learn through guided and recreated exploration. *Educational Studies in Mathematics*, 112(1), 49-72.

Heyd-Metzuyanim, E., & Sfard, A. (2012). Identity struggles in the mathematics classroom: On learning mathematics as an interplay of mathematizing and identifying. *International Journal of Educational Research*, 51-52, 128-145.

Nachlieli, T., & Tabach, M. (2019). Ritual-enabling opportunities-to-learn in mathematics classrooms. *Educational Studies in Mathematics*, 101(2), 253-271.

Lavie, I., Steiner, A., & Sfard, A. (2019). Routines we live by: From ritual to exploration. *Educational Studies in Mathematics*, 101(2), 153-176.

Sfard, A. (2008). *Thinking as communicating: Human development, the growth of discourses, and mathematizing*. Cambridge University Press.

Sfard, A. (2012). Introduction: Developing mathematical discourse—Some insights from communicational research. *International Journal of Educational Research*, 51-52, 1-9.

Sfard, A., & Prusak, A. (2005a). Identity that makes a difference: Substantial learning as closing the gap between actual and designated identities. In Chick, H. L. & Vincent, J. L. (Eds.), *Proc. 29th Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 1, pp. 37-52). PME.

Sfard, A., & Prusak, A. (2005b). Telling identities: In search of an analytic tool for investigating learning as a culturally shaped activity. *Educational Researcher*, 34(4), 14-22.

# PROSPECTIVE MATH TEACHERS' VISION OF HIGH-QUALITY MATHEMATICS INSTRUCTION WITH TECHNOLOGY: A FOCUS ON ROLE OF THE TEACHER

Allison W. McCulloch[1], Lara K. Dick[2], Nina G. Bailey[3], Jennifer N. Lovett[4], Joshua Wilson[4], Purity Muthitu[5]

[1]UNC Charlotte, [2]Bucknell, [3]Montclair State, [4]Middle Tennessee State, [5]NC State

*We report on a study of preservice secondary mathematics teachers' instructional vision evolves as a result of engaging in practice-based approach to learning how to teach mathematics. Findings show that preservice teachers developed a more sophisticated vision of the role of the teacher.*

## INTRODUCTION

Internationally, educational leaders assert the importance of preparing teachers to teach mathematics with technology (e.g., AMTE, 2017; AITSL, 2011). This means understanding the important role a teacher plays when using high-quality technology-enhanced mathematics tasks to support student learning. Much of the research focused on prospective secondary mathematics teachers' (PSTs') preparedness to teach with technology has focused on self-reported confidence in their preparedness (e.g., Karatas et al., 2017), their design, adaptation, and/or selection of technology-enhanced tasks (e.g., Akapame et al., 2019), and the ways they position the technology in a lesson (e.g., Hollebrands et al., 2016). However, from these perspectives little attention is given to the ways that PSTs envision a teacher's role in implementing those tasks effectively. The role of the teacher is emphasized when teacher preparation programs take a practice-based approach (Grossman et al., 2009) to learning to teach mathematics with technology. Since PSTs are still developing their practice, attending to their instructional vision can provide insight into their future practice, including the role of the teacher. One's vision of high-quality mathematics instruction (VHQMI) provides insight into the ways in which one aspires to enact instruction (Hammerness, 2001). The purpose of this paper is to examine how PSTs' VHQMI evolves as a result of engaging in a practice-based approach to learning how to teach mathematics with technology with a focus on the role of the teacher.

## BACKGROUND LITERATURE

In their review of technology-enhanced pedagogy in teacher learning, Zinger et al. (2017) called for less attention to PSTs' use of technological tools and more attention on the role of the teacher in using those tools to address problems of practice, noting that looking at PSTs' pedagogical change over time would provide important insight to their developmental trajectory. However, PSTs are still refining their practice and during their early coursework and often do not have opportunities to plan or implement lessons, particularly lessons that include technology-enhanced mathematics tasks (McCulloch et al., 2021). With this in mind, rather than attending to their enacted



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 209-216). PME.

instruction, researchers have called for attending to PSTs' *instructional vision* as an indication of their developmental progress during teacher preparation programs (e.g., Feiman-Nemser, 2001). Munter (2014) described instructional vision as "ways of seeing the world that encompass horizons not yet reached" (p. 587).

Research has shown that teacher preparation programs focused on pedagogies of practice (Grossman et al., 2009) help shift PSTs' visions of mathematics instruction toward visions of teaching that are less aligned with teaching as telling and more aligned with ambitious teaching practices (e.g., Arbaugh et al., 2021; Jansen et al., 2020; Walkowiak, et al., 2015). However, Jansen et al.'s (2020) study of early career teachers' VHQMI and perceptions of how their teacher preparation program influenced their VHQMI found that technology integration was missing from teachers' vision.

## THEORETICAL FRAMEWORK

Instructional vision is a discourse that teachers (including PSTs) "employ to characterize the kind of ideal classroom practice to which they aspire but have not yet necessarily mastered" (Munter & Wilhelm, 2021 p. 343). As such, one's instructional vision is an expression of their appropriation of the principles, frameworks, and ideas about teaching and learning that they have encountered through their personal and professional learning experiences (Munter & Wilhelm, 2021). Munter (2014) described a specific vision of mathematics instruction deemed "high-quality" that is aligned with the literature on effective mathematics instruction, guiding frameworks in mathematics education (e.g., NCTM, 2014), and data collected middle school classrooms (Cobb & Smith, 2008). Munter's VHQMI framework includes three interrelated rubrics: role of teacher, classroom discourse, and mathematical tasks.

Specific to the role of the teacher, Munter (2014) includes five levels that each describe the extent to which teachers express an instructional vision consistent with VHQMI – teacher as *motivator*, *deliverer of knowledge*, *monitor*, *facilitator*, or *more knowledgeable other*. At Level 0, *motivator*, the primary role of the teacher is to entertain and/or keep students focused on the lesson. Level 1, teacher as *deliverer of knowledge*, is characterized by the teacher being responsible for delivering the mathematical knowledge clearly and accurately. Level 2, *monitor*, is characterized by the teacher demonstrating what students are to be doing and then giving students time to work independently or in small groups with the teacher walking around the room to address questions and/or direct students down a particular solution path. Level 3, *facilitator*, involves the teacher launching a task, students working in groups on the task, and then a whole class discussion where solution strategies are shared. When students are working in groups the teacher will monitor and pose assessing and advancing questions without directing students toward a particular strategy. Level 4, teacher as *more knowledgeable other* (highest level), is characterized by the teacher and students sharing the mathematical authority in the lesson, working together toward a shared mathematical goal, with the teacher proactively supporting student learning through selecting, sequencing and questioning.

## METHODS

11 PSTs who took courses that used curriculum materials specifically designed to prepare PSTs to teach secondary mathematics with technology including activities from the Preparing to Teach Mathematics with Technology – Examining Student Practices [PTMT-ESP] materials (go.ncsu.edu/ptmt) participated in this study. The design of the PTMT-ESP materials draws on the literature on practice-based teacher education (e.g., Grossman et al., 2009; McDonald et al., 2013).

### Data Collection

Data included PSTs' written responses to a pre and post course vision prompt adapted from Munter (2014). It stated: If you were asked to observe a technology-using math teacher's classroom for one or more lessons, what would you look for to decide whether the mathematics instruction (including the use of technology) was high quality? In your response make sure you describe what you would expect to see/hear from the teacher, students, and mathematical tasks during your observations.

### Data Analysis

Data were blinded and then analysed using Munter's (2014) VHQMI rubric levels for *role of teacher*. N/A was coded when a PSTs did not include any information about the role of the teacher. All vision statements were coded by at least 3 members of the research team and discrepancies were discussed until consensus was reached with the entire team. Finally, similar to Walkowiak et al. (2015), the data were organized into three groups, PSTs with a lower vision score at the end of the semester, those with no change in level, and those with an increased score, and analysed for emerging themes.

## FINDINGS

The data are represented in Table 2. At the beginning of the semester all but two of the PSTs' *role of teacher* scores were 2 or lower. This suggests the PSTs gave very little consideration to the important role that teachers play in the facilitation of high-quality technology-enhanced mathematics tasks when they began the course. At the end of the semester, two PSTs decreased in their *role of teacher* score, four stayed the same, and five increased. In the following sections, findings related to each group are described.

| PST # | 2 | 3 | 4 | 5 | 6 | 7 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Pre | 2 | 3 | 2 | 0 | N/A | 2 | 3 | 1 | 2 | 3 | 2 |
| Post | 1 | 3 | 4 | 4 | 4 | 2 | 2 | 3 | 2 | 3 | 4 |
| Level Change | -1 | 0 | 2 | 4 | 5 | 0 | -1 | 2 | 0 | 0 | 2 |

Table 2: Rubric Scores *Role of Teacher*

**Decrease in Role of Teacher Score**

Two PSTs decreased in their *role of teacher* score from the beginning to the end of the semester. The decreases were not large (one level). At the beginning of the semester, PST 2 described the *role of teacher* as a *monitor* (level 2), describing that the teacher walk around the class as students were working and ask assessing questions focused on procedures. PST 2 wrote, “From the teacher, I would most likely hear questions being asked – ‘how did you get that answer?’ ‘Could you show me how?’” At the end of the semester PST 2 used a specific example, a Jeopardy game, to illustrate their vision. In this example, the *role of the teacher* was relegated to “controlling the game pace and flow” and “going over each solution” (level 1).

PST 9’s vision statement at the start and end of the semester were very similar. Both focused on the teacher as *monitor*, but at the beginning of the semester PST 9 included how they would facilitate students’ engagement as they did so (level 3) stating,

> I would expect to see the teacher walking around, helping those students that raise their hand, but not [giving] them the entire answer. Asking them to question themselves and figure it out themselves. In this way the students would productively struggle.

At the end of the semester, PST 9’s vision statement included monitoring with a focus on guiding students through the activity rather than facilitating (level 2). For example,

> The teacher should be watchful for students doing nothing or raising their hands… when it’s pretty clear most have done enough of the task to have learned the lesson, the teacher will check in with them, perhaps asking assessing questions, or performing the task they just completed on the projector, or showing them a completed version so that they know whether they did it correctly.

PST 9’s post vision statement lacked the reference to students’ productive struggle and instead focused on the teacher’s role as controlling how students’ progress.

**No Change in Role of Teacher Score**

Four PSTs (PST 3, 7, 11, 12) had no change in their *role of teacher* rubric score from the beginning to the end of the semester. While there was no change, further analysis did reveal noticeable differences from pre to post that suggest their vision of the *role of the teacher* was moving toward being aligned with the next rubric level.

Both PST 7 and 11 described a *role of teacher* focused on monitoring students as they work on a technology-enhanced math task (level 2). At the start of the semester, they both described monitoring to ensure on-task behaviour. For example, PST 11 wrote, “You could hear the teacher walking around, checking-in on the students. Making sure they were on task and learning.” At the end of the semester both PSTs described more purposeful monitoring. PST 11 wrote that they envisioned “a teacher walking around the room asking for explanations of what students are doing and pushing them to figure out their questions/explanations on their own to solidify their thinking.” While both remained at level 2, their focus shifted from making sure students were behaving as expected to consideration of student thinking while monitoring.

Remaining at level 3, PST 3 and 12 both described the *role of the teacher* as a *facilitator*. They described intentional moves to facilitate student progress on a technology-enhanced task. For example, PST 12 wrote,

> Some things I might hear from the teacher include guided questions to help the students go in a specific direction and questions to the students that force the students to explain themselves and become more confident in their findings rather than just assuming the correct answer.

PST 12 is envisioning the teacher facilitating students' work on a technology-enhanced math task in a specific way (i.e., "a specific direction"). At the end of the semester, PST 12's description of facilitation emphasized understanding student thinking rather than moving students through the task. They wrote, "I would also expect to see the teacher up and walking around the classroom observing what each of the students are doing…I would expect to hear her ask questions that make the students dig a little deeper into their thinking." These PSTs consistently described the *role of teacher* as *facilitator* (level 3), however they moved from focusing on student progress on the task to students' mathematical thinking. All PSTs that maintained the same level demonstrated movement toward the next level.

**Increase in Role of Teacher Score**

Five of the PSTs increased in their *role of teacher* score. Three PSTs moved up two levels, one moved up 4 levels, and one moved up 5 levels. All of the PSTs scored at a level 2 or lower at the beginning of the semester. At the end of the semester, one PST scored at level 3 and the other four scored at level 4.

At the beginning of the semester PST 10 envisioned the *role of the teacher* as a *deliverer of knowledge* (level 1) using an "I do you, we do, you do" lesson structure:

> Then the teacher hands out a guided notes worksheet and assignment worksheet. The students listen to the teacher talk about the content and take notes. At the bottom of the notes worksheets the students are prompted to open up Desmos… Once the class has had adequate time to complete those problems the teacher then instructs them to complete the rest of the problems individually. Once the students have had enough time the class and teacher go over the answers together.

At the end of the semester, PST 10 describes the *role of the teacher* as a *facilitator* (level 3), focused on listening to students and watching how they engage mathematically with the technology in the task. PST 10 wrote,

> I would be looking to see if the teacher was walking around engaging in student conversations. I would expect the teacher to ask questions to further student thinking, listen to what students were discussing when completing the activity, and watching how students interact with the technology to determine solutions…After students worked through the activity, I would expect the teacher to pull multiple different examples from the activity and have a class discussion on what students see and what they wonder.

As another example, PST 4 described the *role of teacher* as *monitor* (level 2), by vaguely noting that after providing a technology-enhanced math task the teacher would "monitor the students during this time." At the end of the semester, PST 4 envisioned the *role of teacher* as *more knowledgeable other* (level 4). Their vision statement included descriptions of the teacher intentionally planning for and proactively supporting student. For example, PST 4 wrote

> With a sufficient high-quality task, an instructor should facilitate a discussion on the topics using the students' work. This is done through carefully cultivating the responses and ordering them for discussion… The instructor guides and facilitates, occasionally may need to step in to reach the necessary daily goals, but should almost never directly explain a topic in full without student input. Also, the teacher should, when students are exploring and completing the task, ask students assessing questions that help the teacher see students' understanding then ask advancing questions that lead students to a more developed understanding. This monitoring of students should be done intermittently and with anticipated outcomes in order to hasten the instruction but also make time for the instructor to plan how to facilitate any whole class discussions. Lastly, behind each lesson, the instructor should anticipate and plan instruction, but arguably more importantly reflect on their instruction to improve it for the future and potentially modify and plans to make sure students reach the desired learning goals for a given time period.

PST 4's vision statement was similar to the other PSTs that increased in level. It highlighted pedagogical practices such as an effective task launch, anticipating student thinking, planning assessing and advancing questions based on anticipations, creating a monitoring chart, and planning for facilitating a whole class discussion. In each of these descriptions of their VHQMI the PSTs envisioned the role of the teacher in a way that is consistent with the literature on effective teaching using technology-enhanced math tasks. What set them apart is the attention to student thinking both in the planning and enactment of the envisioned instruction.

## DISCUSSION AND CONCLUSION

Overall, most of the PSTs in this study had a more sophisticated vision of the role of the teacher (including those with no change in their rubric scores) when teaching secondary mathematics using technology at the end of the semester. This is consistent with other work that included an intervention that is practiced-based and aligned with ambitious instruction (i.e., coursework, professional development; e.g., Arbaugh et al., 2021; Jansen et al., 2020; Munter, 2014; Walkowiak et al., 2015). It is promising that specifically focusing on teaching using technology-enhanced tasks, a context that many researchers have shown PSTs do not feel prepared to do (e.g., Wang et al., 2018), resulted in such visions regarding the role of the teacher. This suggests that a practice-based approach to learning to teach mathematics with technology highlights the important role that teachers must play when using technology-enhanced math tasks and possibly contradicts the somewhat common belief of the "technology teaches" (i.e., a technocentrist view) that some PSTs hold (Zinger et al., 2017).

There are several possible explanations for those PSTs that did not increase in rubric level. First, the PSTs that decreased and those that stayed the same all began the semester at a level 2 or 3. Meaning they were already thinking about the ways in which a teacher must attend to students, not just deliver content or entertain (levels 0-1). It may take more than one semester for some PSTs to shift to considering the role of the teacher as a facilitator or more knowledgeable other. Those PSTs that began the course with lower rubric level scores (N/A-1) made larger strides. This may be because the approximations of practice they engaged in were so different from their initial vision, which may have incited deep reflection on their ideas about the role of the teacher when using technology-enhanced math tasks. This is an area in need of further research.

## Acknowledgements

This work was supported by the National Science Foundation [DUE 1820998, DUE 1821054, DUE 1820967, DUE 1820976]. Any opinions, findings, and conclusions or recommendations expressed herein are those of the principal investigators and do not necessarily reflect the views of the NSF.

## References

Akapame, R., Burroughs, E., & Arnold, E. (2019). A clash between knowledge and practice: A case study in TPACK in three pre-service secondary mathematics teachers. *Journal of Technology and Teacher Education*, *27*(3), 269–304. https://www.learntechlib.org/primary/p/208634/

Arbaugh, F., Graysay, D., Freeburn, B. & Konuk, N. (2021). Investigating secondary mathematics preservice teachers' instructional vision: Learning to teach through pedagogies of practice. *Journal of Teacher Education, 72*(4), 448–461. https://doi.org/10.1177/0022487120965938

Association of Mathematics Teacher Educators [AMTE]. (2017). *Standards for preparing teachers of mathematics*. https://amte.net/standards

Australian Institute for Teaching and School Leadership [AITSL]. (2011). *Australian professional standards for teachers*. https://www.aitsl.edu.au/standards

Cobb, P. & Smith, T. (2008). District development as a means of improving mathematics teaching and learning at scale. In K. Krainer & T. Wood (Eds.), *International handbook of mathematics teacher education: Vol. 3. Participants in mathematics teacher education: Individuals, teams, communities, and networks* (pp. 231–254). Sense Publishers.

Feiman-Nemser, S. (2001). From preparation to practice: Designing a continuum to strengthen and sustain teaching. *Teachers College Record*, *103*(6), 1013–1055. https://doi.org/10.1111/0161-4681.00141

Grossman, P., Hammerness, K., & McDonald, M. (2009). Redefining teaching, re-imagining teacher education. *Teachers and Teaching: Theory and Practice*, *15*, 273–289. https://doi.org/10.1080/13540600902875340

Hammerness, K. (2001). Teacher's visions: The role of personal ideals in school reform. *Journal of Educational Change, 2*(2), 143–163. https://doi.org/10.1023/A:101791615264

Hollebrands, K. F., McCulloch, A. W., & Lee, H. S. (2016). Prospective teachers' incorporation of technology in mathematics lesson plans. In Niess, M., Driskell, S., & Hollebrands, K. (Eds)., *Handbook of Research on Transforming Mathematics Teacher Education in the Digital Age* (pp.272–292). IGI Global.

Jansen, A., Gallivan, H. R., Miller, E. (2020). Early-career teaches' instructional visions for mathematics teaching: impact of elementary teacher education, *Journal of Mathematics Teacher Education*, 23, 183–207. https://doi.org/10.1007/s10857-018-9419-1

Karatas, I., Tune, M. P., Yilmaz, N., & Karci, G. (2017). An investigation of technological pedagogical content knowledge, self-confidence, and perception of pre-service middle school mathematics teachers towards instructional technologies. *Journal of Educational Technology & Society*, *20*(3), 122–132. https://www.jstor.org/stable/26196124

McCulloch, A. W., Leatham, K. R., Lovett, J. N., Bailey, N. G., & Reed, S. D. (2021). How we are preparing secondary mathematics teachers to teach with technology: Findings from a nationwide study. *Journal for Research in Mathematics Education, 52*(1), 94–107. https://doi.org/10.5951/jresematheduc-2020-0205

McDonald, M., Kazemi, E., & Kavanagh, S. S. (2013). Core practices and pedagogies of teacher education: A call for a common language and collective activity. *Journal of Teacher Education*, *64*, 278–386. https://doi.org/10.1177/0022487113493807

Munter, C. (2014). Developing visions of high-quality mathematics instruction. *Journal for Research in Mathematics Education, 45*(5), 584–635. https://doi.org/10.5951/jresematheduc.45.5.0584

Munter, C. & Wilhelm, A.G. (2021). Mathematics teachers' knowledge, networks, practice, and change in instructional visions. *Journal of Teacher Education*. *72*(3), 342–354. https://doi.org/10.1177/0022487120949836

National Council of Teachers of Mathematics. (2014). *Principles to actions: Ensuring mathematical success for all*. Author.

Walkowiak, T. A., Lee, C. A., & Whitehead, A. (2015). The development of preservice teachers' visions of mathematics instruction. Paper presented at the *Research Conference of the National Council of Teachers of Mathematics*, Boston, MA.

Wang, W., Schmidt-Crawford, D., & Jin, Y. (2018). Preservice teachers' TPACK development: A review of literature. Journal of Digital Learning in Teacher Education, *34*(4), 234–258. https://doi.org/10.1080/21532974.2018.1498039

Zinger, D., Tate, T. & Warschauer, M. (2017). Learning and teaching with technology: Technological pedagogy and teacher practice. In D. J. Clandinin & J. Husu (Eds.), *The SAGE Handbook of Research on Teacher Education* (pp. 577–593). SAGE.

# LEARNING TO TEACH MATHEMATICS WITH INSTRUCTIONAL TECHNOLOGY: A PRAXEOLOGICAL ANALYSIS OF A SWEDISH MATHEMATICS TEACHER EDUCATION COURSE

Farouq Sessah Mensah

Department of Teaching and Learning, Stockholm University

*The study used the Anthropological Theory of the Didactic to understand how, why, and what is privileged in a Swedish mathematics teacher preparation course. A mathematics teacher-educator interview was analysed using a reference model. The study results show that teaching with instructional technology in cognitive ways was privileged pre-didactically, didactically, and post-didactically. The didactical cognitive praxis addressed teachers' and learners' use of instructional technologies. The privileged didactical logos linking the praxis were competencies (förmågor) in the Swedish upper secondary curriculum. The privileged meta-didactical praxis was the decomposition, representation and approximation of practice, with some implicit meta-didactical logos discussed.*

## INTRODUCTION

Mathematics education discussions have long anticipated that integrating instructional technologies into core teaching practices will lead to, or at least have the potential to lead to, shifts in both teaching and learning as well as content and context (e.g., Hillmayr et al., 2020). This study uses the term instructional technologies to refer to the integration of digital technologies into the classroom to enhance productivity and transform the depth of knowledge acquisition among teachers and students. I will refer to instructional technologies as 'IT' hereof. IT cannot simply be relied upon to realise its full potential. Instead, a careful and deliberate didactical approach is necessary to realise its potential and minimise adverse effects. However, research findings indicate that while teachers recognise the importance of IT as a teaching and learning tool and are eager to use it, they are often challenged to make it an integral part of their teaching (e.g., Valtonen et al., 2020).

Research has shown that the quality of teacher education (TE) programmes significantly impacts pre-service teachers' (PTs') ability to integrate IT meaningfully into their teaching (e.g., Kafyulilo & Fisser, 2019). Although efforts by TE programmes on IT integration have been beneficial, they have been criticised for failing to adequately provide PTs with the hands-on experience to integrate IT into their future teaching (Fathi & Ebadi, 2020). Consequently, it is essential to understand how, why, and what is privileged in mathematics teacher preparation courses that engage PTs in learning to teach mathematics with IT. As a starting point, I explore a case of a Swedish mathematics teacher preparation course.



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 217-224). PME.

## THE SWEDISH CASE

The Swedish National Agency for Education (Skolverket, 2010) has criticised Swedish schools for their low integration of IT into classroom practices, stating that they have IT-equipped teaching-learning environments but struggle to develop and evaluate innovative teaching strategies that integrate IT. Different sources, including the European Commission, the Swedish government, and researchers, have acknowledged this issue. How Swedish TE programmes prepare PTs to integrate IT into their core teaching practices remains unclear. Moreover, evidence suggests limited research on the Swedish mathematics TE and, more specifically, IT trends (e.g., Dewi et al., 2021).

## THEORETICAL CONSIDERATION AND RESEARCH QUESTIONS

The study is grounded in the Anthropological Theory of the Didactic (ATD), which considers doing, teaching, diffusing, creating, and transposing mathematics, as well as any other kind of knowledge, as human activities taking place in institutional settings (Chevallard & Bosch, 2020). ATD indicates that any human activity can be understood as a praxeology. A praxeology has two constituent parts: the praxis and the logos. Tasks that need to be solved and techniques for doing so constitute the praxis. In contrast, logos constitutes technology, serving as a discourse for the techniques and theory which justifies the technological discourse. ATD discusses two types of praxeologies: mathematical praxeologies (MP) and didactic praxeologies (DP) (Chevallard & Bosch, 2020). A MP is when the types of tasks, techniques, technologies and theories are mathematical.

Conversely, DP is when the types of tasks, techniques, technologies and theories support the teaching or learning of a MP. DP generally refers to the activities of teachers in schools. However, in TE, the DP is the knowledge at stake, as PTs are taught how to teach. Further, Arzarello et al. (2014) discuss meta-didactical praxeologies, which refer to the tasks, techniques, and justifying discourses that emerge during the TE process where the teaching of mathematics is the knowledge at stake. From a mathematics TE perspective, this is a DP with ‘mathematics education’ to be taught and learned. However, for clarity, I will use the term meta-didactical praxeology when didactical knowledge is at stake.

From a praxeological perspective, the following research questions guided the study.

1. What didactic praxeologies for integrating IT into mathematics teaching are privileged by a mathematics teacher educator?
2. How are meta-didactical praxeologies privileged when a mathematics teacher educator describes her teaching of IT integration in mathematics teaching?

## RESEARCH METHOD

The study is designed as a single case study grounded in ATD. This case serves as a preliminary analysis to provide a nuanced, empirically rich, context-specific account of a more extensive study that will include multiple cases. The study design allows for an initial critical reflection, analysis, interpretation, and understanding of the research problem.

**Case Context**

The study was conducted in a mathematics teacher preparation course for upper secondary school PTs, focusing on mathematics didactics (methods/education) and field placement (verksamhetsförlagd utbildning, VFU). The module focused on mathematics teaching and learning following the current regulatory documents for upper secondary schools and pertinent research in the didactics of mathematics. The study involved 26 PTs and a mathematics teacher educator (MTE) from a Swedish university. Ofelia (pseudonym), the MTE, was the principal participant. The course was selected based on engagement with IT and the willingness of the PTs and MTE.

**Data Collection and Analysis of Case**

Data for the study was collected from a semi-structured interview with a MTE. The interview was conducted to understand the course structure, the privileged didactical and the meta-didactical praxeologies of IT in the teacher preparation course. The interview was audio recorded and transcribed. The interview transcripts were analysed by drawing on a reference model (RM) and using software for qualitative data analysis.

**Reference Model**

A reference model is a framework researchers use to challenge and address didactical situations in specific institutions (see Wijayanti & Winslow, 2017). The RM functions as the analytical framework for the study. To answer research question one, I differentiate between three types of praxis – didactical praxis (activities taking place in the teaching-learning situation), pre-didactical praxis (activities preceding the teaching) and post-didactical praxis (activities following the teaching). I further drew on the categorisations of the use of IT provided by Akapame et al. (2019) and Clark-Wilson et al. (2020). Akapame et al. (2019) categorise IT use as cognitive or productive. Cognitive use involves using IT to develop conceptual understanding, which opens new forms of engagement with the content, while productive use maximises users' output by applying various technological means.

Clark-Wilson et al. (2020) classify IT uses into four different categories: two are about pre- and post-didactical praxeologies (planning or keeping track of assessment results, collaboration with other teachers), and two are didactical, namely representing and learners' independent work. Further, I made a distinction according to who would be the agent in using the IT: the teacher or the learner. In this praxeological analysis, I could not clearly distinguish between the type of task and the technique for carrying out the task. Hence, I put them together as the praxis in the RM for the didactical knowledge at stake. To answer research question two, I drew on Grossman et al. (2009) practice-based TE of representation, decomposition, and approximation of practice as RM to explore the meta-didactical praxis the MTE employs. Representation is when PTs work with IT as part of their learning process in the TE programme or observe the

MTE using IT. Decomposition is when PTs analyse what works and does not work using different IT in different task situations. The approximation is when PTs enact mathematics teaching with IT and reflect on their experience. The logos component for the didactical and meta-didactical praxis was determined inductively by looking for justifications for the identified praxis components.

## RESULTS

### Didactical Praxeology

Generally, a didactical praxeology of using IT to teach and learn mathematics in a cognitive way was privileged. Pre-didactical cognitive praxis was privileged in the application of variation theory to a given task, analysing the mathematics involved in the task, reading educational text and writing assignments, setting teaching goals and connecting them to förmågor (competencies), and distinguishing between GeoGebra tasks with high and low cognitive demands. Ofelia privileged pre-didactical cognitive praxis, describing what was covered in the seminar and giving examples of tasks of high and low cognitive demands. The following are excerpts on pre-didactical cognitive praxis privileged.

> So, the whole idea for me in this seminar was to talk about the cognitive demands of this task, the mathematics and questions about variation theory. […] creating a pentagon by hand and […] calculating the area, […] you must think a bit. Maybe connect different procedures or so. So, we discussed that maybe this would be level two or three on cognitive demands, but by doing it in GeoGebra, it is like one click, and it is about zero on cognitive demands. So, that is one way GeoGebra reduces cognitive demand to zero.

Ofelia describes that while IT has the potential to increase cognitive demands, it can only do so if it is used effectively and meaningfully, which I interpreted as a privileging of cognitive praxis due to the emphasis on cognitive demands. Ofelia referenced a problem from Granberg and Olsson (2015, p. 53), which PTs were privileged to as a 'true problem'. The instructions for the PTs and Ofelia's comments about the task were as follows:

> (1) Try to complete the GeoGebra task the study is about. (2) The authors distinguish between 'imitative reasoning' and 'creative reasoning'. Describe what they mean, give examples from your work with the task, and based on the article. (3) The authors believe that GeoGebra can constitute a didactic environment by providing creative feedback on student work, even though the only feedback is that the lines to given functions are shown. Exemplify and discuss the claim.

> […] they solve a problem where they have four lines and change the equations to tilt the square. Moreover, when you ask them to do it together, they get to reason and try to figure out different ways of doing this. You saw the guy who animated it to turn around and move. You must consider the kind of tasks and questions you ask when introducing GeoGebra.

In this task, Ofelia describes how the task keeps the cognitive demand relatively high and enables the students to reason with GeoGebra. This practical demonstration privileged a first-hand experience of how IT functions and learners' interactions with it. Ofelia reiterated as part of the pre-didactical cognitive praxis the importance of PTs being mindful of the tasks and questions they introduce when implementing any IT.

The didactical cognitive praxis privileged in teacher and learner use focused on how PTs could use IT to increase the depth with which teachers and students can acquire new knowledge. The excerpt below shows what was privileged as a didactical cognitive teacher and learner use of IT.

> […] what we try to do in general is model that kind of teaching, where we try to show them how this can be used and what they can do, and as the next step, how they can do this with learners.

A privileging of didactical cognitive praxis in teacher and learner use was seen when Ofelia discussed the potential for PTs to utilise IT not only for their own instruction but also for students' use during the learning process.

A privileging of post-didactical cognitive praxis was seen when Ofelia discussed examples of PTs discussing their experiences from their field placement and making connections to the förmågor, exemplified in the following excerpt.

> […] they have experience teaching for three weeks, and now we can start to talk about what we can do with digital tools in such a setting. They have done so much in a practicum that we had about 1 hour to summarise. So, what I asked them to do during this hour was to bring one or two of their lesson plans to the groups and discuss them together. How can the lesson plan connect to the förmågor?

A privileging of post-didactic cognitive praxis was also seen when Ofelia discussed the importance of connecting different parts of lesson plans to the förmågor in post-lesson discussions. For example, if the emphasis was on problem-solving, the PTs discussed how it is shown in the lesson plan under consideration.

In a few excerpts, there was a privileging of productive praxis. For example, Ofelia described how GeoGebra saves time in the following excerpt.

> […] to show them that this can be used for demonstrations, more or less as I did. I […] had time to move into the different parts of GeoGebra and show when you use it […] to make your own constructions.

This is an instance where PTs were privileged in Ofelia's descriptions to learn to use IT to increase the efficiency with which they can accomplish their tasks.

Ofelia privileged the didactical logos linking the praxis as förmågor, which aligns with the Swedish upper secondary school curriculum. The excerpt below shows the didactical logos privileged in Ofelia's descriptions.

> We provide them with tools such as mathematical competencies, which are close to the Swedish curriculum so that they can understand the goals and the aims of the mathematics teaching curriculum.

This excerpt provides a justification for the didactical praxis. The PTs were privileged in Ofelia's descriptions to become familiar with the work of a mathematics teacher and, at the same time, gain experience in planning and leading teaching.

## Meta-Didactical Praxeology

The most described meta-didactical praxis privileged in Ofelia's discussion of teaching PTs how to teach mathematics using IT was the decomposition of practice, where PTs discuss, reflect on and analyse different task situations in the course. As an example of the privileged practice of decomposition, I refer to Granberg and Olsson's (2015, p. 53) task given in the course. In this task, PTs were required to analyse different types of mathematical reasoning and examine the didactic environment using theoretical concepts. Ofelia also privileged representation of practice, describing how she modelled the use of IT to teach mathematics, as in the following excerpt.

> […] to show them that this can be used for demonstrations, more or less as I did. I […] had time to move into the different parts of GeoGebra and show when you use it […] to make your own constructions.

Demonstrations and showing PTs how to use IT to teach mathematics are examples of representation interpreted as a privileging of representation as a meta-didactical praxis, mainly in demonstrations. However, approximation of practice was a rarely privileged meta-didactical praxis in the course. For example, Ofelia described the extent to the approximation of practice in the following excerpt.

> That would be when they present the tasks. That is how far we come in teaching practice besides the practicum. So, they have the entire three weeks of practicum where you saw that it was like 90% who had done something digital in my group. So, it is prevalent. However, the teaching practice [in the course] is more like they present the task, and we ask them to plan something interactively for their peers.

Other than the practicum, PTs were privileged to approximate practice through presentations. The privileged meta-didactical logos was the decision that students need to learn about teaching in general and get a sense of different ways of doing a task before learning about teaching with IT. Ofelia described her meta-didactical logos in the following excerpt.

> We talked a lot about having the digital tools before practicum. However, for novice teachers, we decided in the end […]. To prepare them for teaching for the first time, we decided that perhaps this is not the best time for digital tools. So, we do something with what do you say? It focuses less on assessment and more on having a good discussion and learning opportunities.

Although implicit, this example justifies what was privileged in Ofelia's descriptions of her meta-didactical praxis. Ofelia privileged learning general teaching concepts before IT integration and how PTs could orchestrate productive mathematics discussions.

## CONCLUDING DISCUSSIONS

Using the RM, the paper exemplifies how a combination of theoretical constructs can be operationalised to study how teaching practices are privileged in TE. The single case data shows that PTs are privileged to learn to teach mathematics with IT in cognitive ways. This result is consistent with the visions presented in the previous literature (e.g., Akapame et al., 2019), which propose that PT preparation programmes should include teaching PTs to use IT in cognitive ways.

In Ofelia's discussion, I interpreted a privileging of pre-didactical, didactical, and post-didactical learning opportunities and practise teaching mathematics with IT, maximising the course's use of IT as a cognitive resource. PTs learned how to teach mathematics using IT for their cognitive development and that of their learners within the didactical context. This result is consistent with much of the previous research, which indicates that mathematics TE courses on teaching how to teach using IT often cover at least two of the three types of praxis distinguished in the RM (e.g., Kafyulilo & Fisser, 2019).

Although used differently and at different phases in the course, the meta-didactical praxis of decomposition, representation, and approximation overlap and reinforce each other. On the assumption that the preparation programme is in its early stages, the study's findings showed that the decomposition of practice was mainly privileged. The result is consistent with Herbst et al. (2019), who engaged novice teachers in learning experiences that modelled practice elements and allowed student teachers to deconstruct, critique, and reflect on the essential components of teaching mathematics with IT. While there have been requests for more holistic approaches to teaching and learning complicated practices, Grossman et al. (2009) argued that breaking down the practice into manageable chunks and working on each individually before putting it all together can be highly effective. This appears to be privileged in the case reported here. Using theoretical concepts from mathematics education becomes one way to link logos and praxis, theory and practice, supporting PTs more strongly in developing professional orientation. Though Ofelia could substantiate the choice of sequencing in the course, the explicit reference to meta-didactical logos was limited. This raises questions about the extent to which mathematics TE can be further developed through a more robust grounding in research.

**Acknowledgement**

I acknowledge the REMATH research school, funded by the Swedish Research Council (2021-00534) and Helge Ax:son Johnsons stiftelse (Nr F23-0012).

**References**

Akapame, R., Burroughs, E., & Arnold, E. (2019). A clash between knowledge and practice: A case study of TPACK in Three pre-service secondary mathematics teachers. *Journal of Technology and Teacher Education, 27*(3), 269–304.

Arzarello, F., Robutti, O., Sabena, C., Cusi, A., Garuti, R., Malara, N., & Martignone, F. (2014). Meta-didactical transposition: A Theoretical model for teacher education programmes. In A. Clark-Wilson, O. Robutti, & N. Sinclair (Eds.), *The Mathematics Teacher in the Digital Era: An International Perspective on Technology Focused Professional Development* (pp. 347–372). Springer Netherlands. https://doi.org/10.1007/978-94-007-4638-1_15

Chevallard, Y., & Bosch, M. (2020). Anthropological theory of the didactic (ATD). I S. Lerman (Red.). In *Encyclopedia of Mathematics Education* (pp. 53–61). Springer International Publishing. https://doi.org/10.1007/978-3-030-15789-0

Clark-Wilson, A., Robutti, O., & Thomas, M. (2020). Teaching with digital technology. *ZDM, 52*(7), 1223–1242. https://doi.org/10.1007/s11858-020-01196-0

Dewi, N. R., Rusilowati, A., Saptono, S., Haryani, S., Wiyanto, W., Ridlo, S., Listiaji, P., & Atunnisa, R. (2021). Technological, pedagogical, content knowledge (TPACK) research trends: A systematic literature review of publications between 2010-2020. *Journal of Turkish Science Education, 18*(4), 589–604.

Fathi, J., & Ebadi, S. (2020). Exploring EFL pre-service teachers' adoption of technology in a CALL program: Obstacles, motivators, and maintenance. *Education and Information Technologies, 25*(5), 3897–3917. https://doi.org/10.1007/s10639-020-10146-y

Granberg, C., & Olsson, J. (2015). ICT-supported problem solving and collaborative creative reasoning: Exploring linear functions using dynamic mathematics software. *The Journal of Mathematical Behaviour, 37*, 48–62. https://doi.org/10.1016/j.jmathb.2014.11.001

Grossman, P., Compton, C., Igra, D., Ronfeldt, M., Shahan, E., & Williamson, P. W. (2009). Teaching practice: A cross-professional perspective. *Teachers College Record: The Voice of Scholarship in Education, 111*(9), 2055–2100. https://doi.org/10.1177/016146810911100905

Herbst, P., Chazan, D., Chieu, V. M., Milewski, A., Kosko, K. W., & Aaron, W. R. (2019). Technology-mediated mathematics teacher development: Research on digital pedagogies of practice. In *Pre-Service and In-Service Teacher Education: Concepts, Methodologies, Tools, and Applications* (pp. 194–222). IGI Global. https://doi.org/10.4018/978-1-5225-7305-0.ch010

Hillmayr, D., Ziernwald, L., Reinhold, F., Hofer, S. I., & Reiss, K. M. (2020). The potential of digital tools to enhance mathematics and science learning in secondary schools: A context-specific meta-analysis. *Computers & Education, 153*, 103897. https://doi.org/10.1016/j.compedu.2020.103897

Kafyulilo, A., & Fisser, P. (2019). Developing TPACK in science and mathematics teacher education in Tanzania: A proof-of-concept study. In J. Pieters, J. Voogt, & N. Pareja Roblin (Eds.), *Collaborative Curriculum Design for Sustainable Innovation and Teacher Learning* (pp. 139–155). Springer International Publishing. https://doi.org/10.1007/978-3-030-20062-6_8

Skolverket. (2010). *Redovisning av uppdrag om uppföljning av IT-användning och IT-kompetens i förskola, skola och vuxenutbildning* (Dnr U2007/7921/SAM/G). https://www.skolverket.se/getFile?file=2373

Valtonen, T., Leppänen, U., Hyypiä, M., Sointu, E., Smits, A., & Tondeur, J. (2020). Fresh perspectives on TPACK: Pre-service teachers' own appraisal of their challenging and confident TPACK areas. *Education and Information Technologies, 25*(4), 2823–2842. https://doi.org/10.1007/s10639-019-10092-4

Wijayanti, D., & Winslow, C. (2017). Mathematical practice in textbook analysis: Praxeological reference models, the case of proportion. *REDIMAT, 6*(3), 307–330.

# MODELLING WITH EXPERIMENTS – STUDENTS' TRAIT VALUES MEDIATED BY STUDENTS' STATE VALUES

Marielena Menzel[1], Michael Jonscher[2], Stefanie Rach[2], Sebastian Geisler[1]

[1]University of Potsdam, [2]Otto-von-Guericke Universität Magdeburg

*Students' motivation is crucial for successful learning. This study focusses on a repetitive structure of modelling tasks with experiments to examine the development of students' motivation, distinguishing between their stable trait value regarding mathematics in general and their variable state value regarding a certain task. Studies show a tendency for students to dislike modelling tasks. Thus, we chose that context to foster students' motivation. In this quantitative semi-experimental study, 111 secondary school students work on modelling tasks related to linear and exponential functions. Mediation analysis indicate that students' trait and state values are clearly related and that the students' state values partly mediate the changes of students' trait values before and after working on the modelling tasks.*

## INTRODUCTION

A high learning motivation is assumed to be important for successful learning processes (not only) in mathematics. According to expectancy-value theory (Eccles & Wigfield, 2020), the subjective value associated with the task directly influence students' motivation and effort. Rather stable trait values regarding a domain like mathematics can be distinguished from fluctuating state values regarding a specific learning situation or a specific task. Furthermore, within the DYNAMICS-framework, Moeller et al. (2022) propose that state values would mediate the development of trait values. Thus, a beneficial development of students' trait values regarding mathematics could be fostered via positive state values in mathematical learning situations.

We analyse this assumption for the case of mathematics and use the particular important context modelling. Despite the fact that modelling is a central competence in mathematics (OECD, 2017) and has high importance for a modern society (Niss, 1994), many students value modelling less than other mathematics tasks. Thus, more insights in the dynamics behind the development of trait values in the context of modelling are necessary.

## THEORETICAL BACKGROUND

### Expectancy-value theory and state and trait motivation

Following Eccles and Wigfield (2020), the situated expectancy-value theory combines persons' expectancy for success in a certain task and their situated task values to explain students' choices and achievement. Task values comprise four components: *attainment value*, *intrinsic value*, *utility value*, and *costs*. Attainment value describes how important a task is on a personal level but also contains the aspect of how important it is for a person to be good at the task. Intrinsic value describes the



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 225-232). PME.

anticipated enjoyment while doing the task. How relevant a person finds the task for example for the daily life, the career, or upcoming events like exams, is covered by the utility value. The construct of costs reflects the negative aspects of a task. Eccles and Wigfield (2020) differentiate between *effort costs*, describing how much effort working on the task entails in relation to its benefits, *opportunity costs*, describing how working on the task minimizes time or capacity to do other valued tasks, and *emotional costs*, which describe the emotional, psychological, and social costs of possible anxiety or failure. A persons' motivation and intention to work on a certain task is related to the total value derived from these four components (Wigfield & Eccles, 2000).

Early studies that are based on the expectancy-value-theory focused on task values as rather stable dispositions and thus measured general values regarding mathematics (Eccles & Wigfield, 2023). More recent studies differentiate between rather stable trait values that are related to for example mathematics in general and rather situational state values that are related to a specific learning situation or task. These studies show that state values indeed differ between different learning situations (Gaspard et al., 2015). Summing up these studies, Moeller et al. (2022) argue that state and trait values would influence each other. In their DYNAMICS framework, they discuss reciprocal influence between trait values and state values as bottom-up and top-down causalities: bottom-up causalities describe the influence of (repeated) experiences of state values (in similar situations) that could develop into a personal trait value (regarding the topic or activity related to these situations). Top-down causalities refer to the influence of the personal trait values on situational state values. In the current study, we use the terms trait value, to describe how students value their rather stable perspective on mathematics in general, and state value, describing how students value a certain situation.

## Mathematical modelling with experiments and its value for students

Mathematical modelling is a process that takes place when a real-world problem is solved using mathematics (Blum & Leiß, 2007). In the current study, we follow the conceptualisation of an idealized modelling process by Blum and Leiß (2007): the real-world problem has to be understood, simplified, and translated into a mathematical model. The mathematical problem has to be solved and its result retranslated into the real-world. The interpretation and the whole modelling process must be validated. The students' perspective on modelling tasks does not reflect the relevance of that mathematical area. Krawitz and Schukajlow (2018) compared students' values for modelling tasks and values for dressed-up and intra-mathematical tasks and found that students attribute less value to modelling tasks and they considered them less important compared to other types of tasks.

To foster students' motivation for modelling tasks a commonly discussed approach is to combine these tasks with experiments. By involving experiments in modelling tasks, students' motivation can be influenced positively because of the physical activation besides the cognitive stimulation, and the relevance for the reality the students live in (see Ganter, 2013). Studies indicate that modelling tasks connected to experiments can increase students motivation (e.g., Beumann, 2016). In the current study, we understand experiments as a planned and controlled action guided by a hypothesis, for the purpose of gaining knowledge through observation (Ludwig & Oldenburg, 2007). By modelling with experiments, we understand a modelling task that includes a hands-on experiment. The data conducted in the experiment is used for the modelling task.

## THE CURRENT STUDY

The current study is part of the project *Experiments to foster modelling competences and motivation in mathematics* (Ex2MoMa). This study was designed to examine the development of students' trait values during an intervention comprising modelling tasks with experiments. Moeller et al. (2022) used their framework to investigate the influence of students' state values of one learning situation on the next one and on the development of trait values. So far, the framework was used for a study about university students' trait values and state values in lectures. It was not yet applied in other contexts. In the current study, we use the framework in mathematics in the important context of modelling with experiments and analyze the development of trait values in the aforementioned modelling intervention and the role of state values for this development. In particular, we want to answer the following research questions:

1. How do students' trait values develop during the intervention?

   *Hypothesis 1: The values increase and costs decrease as modelling tasks combined with experiments have a motivating effect (see Beumann, 2016).*

2. In which way do students' trait values at the beginning of the intervention affect students' trait values at the end of the intervention (total effects)?

   *Hypothesis 2: All components of trait values (attainment, intrinsic, utility, and costs) at the end of interventions are influenced by the trait values at the begin are significantly positive as trait values are rather stable (Hannula, 2012).*

3. To what extent do students' state values mediate the relationships between the trait values at the beginning of the intervention and the trait values at the end of the intervention (direct and indirect effects)?

   *Hypothesis 3.a: Regarding all components (attainment value, intrinsic value, utility value, and costs) the trait value in the beginning predicts the corresponding state value positively as top-down causality (see Moeller et al., 2022).*

   *Hypothesis 3.b: For all components, the state value predicts the corresponding trait value at the end positively as bottom-up causality (see Moeller et al., 2022).*

   *Hypothesis 3.c: For all components the state value mediates the effect (indirect effect), as a consequence of the hypotheses 3.a and 3.b, but there will be a direct effect left due to the stable character of trait values (see Moeller et al., 2022).*

## METHODOLOGY

### Sample & design

A total of 111 upper secondary students (Age 14 – 18, 48.2 % female) from five classes participated in this study. It took place in three 90-minutes math lessons. In each lesson students worked on an experiment followed by a modelling task. The three tasks differed in the used context (candle burning, stale beer, cooling of tea) but were equally structured, beginning with a situation leading to a hypothesis, the hands-on experiment, and the prompt to model the measured data using functions.

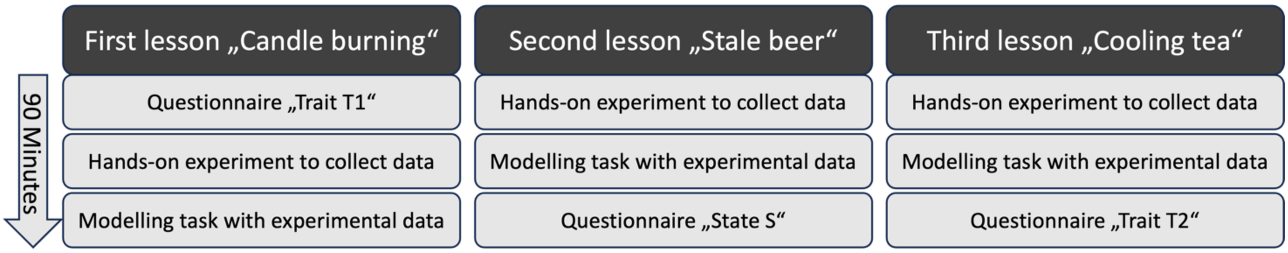



Figure 1: Design of the current study

The modelling task “candle burning” can be solved by using linear functions, the tasks “stale beer“ and „cooling tea” by using exponential functions. Students filled out the identical questionnaires “Trait T1” and “Trait T2” at the beginning and the end of the intervention (see figure 1). They contain statements about students’ trait values regarding mathematics in general. Another questionnaire “State S" is filled out after the students worked on the second modelling task and contains statements concerning state values regarding the task they just worked on.

### Instruments

To measure trait values and state values, we used in each case four approved and slightly adapted scales (Dietrich et al., 2019; Gaspard et al., 2015). All scales of both questionnaires were answered on a six-point likert-scale (*1=not at all true*, *6=totally true*) with good or acceptable reliabilities (T1 Cronbach’s $\alpha > .82$, T2 Cronbach’s $\alpha > .71$, S Cronbach’s $\alpha > .72$). The scales capture the described aspects of each component of the subjective task value: attainment value (AV), intrinsic value (IV), utility value (UV), and costs (CO). Table 1 gives an overview of the used scales. By answering the second and third research questions, we want to explain the development of students’ task values regarding the mechanics of the DYNAMICS framework. To investigate the development of the trait values, we applied a t-test for paired samples. There were 85 samples that could be paired. We used a mediation analysis in Mplus8 (Version 1.8.7) to answer the second and third research question.

| | Trait | | State | |
|---|---|---|---|---|
| Items | # | Example | # | Example |
| AV | 6 | It is important to me to be good at mathematics. | 4 | The task is important to me personally. |
| IV | 4 | I simply like mathematics. | 4 | I enjoyed working on this task. |
| UV | 9 | Mathematics proves to be useful in everyday life. | 4 | Being able to solve such tasks is helpful for my life. |
| CO | 6 | Mathematics is a real burden for me. | 4 | Working on the task was exhausting. |

Table 1: Scales used in the questionnaires, adapted from Dietrich et al. (2019) and Gaspard et al. (2015)

The assumed model for that mediation (for each component of the task value) is that the state value (S) works as a mediator for the relation between the trait value at the beginning of the intervention (T1) and the trait value at the end of the intervention (T2). For the mediation 94 samples were used, as we collected that amount of filled out T1 questionnaires. Missing data at S and T2 were handled using FIML. We calculated four models each for one component (see figure 2, for attainment value). The models for the other components are structured analogous.

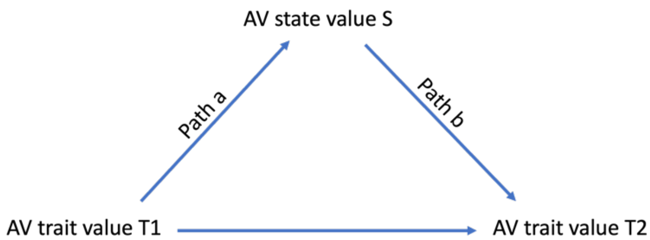



Figure 2: Model for mediation of attainment value

## RESULTS

The descriptive data for all four scales attainment value (AV), intrinsic value (IV), utility value (UV), and costs (CO) in the beginning (T1) and in the end (T2) of the intervention are visible in table 2. In order to answer the first research question, we report the following results of the paired t-tests: Attainment value decreases with a medium effect ($d$ = 0.53) and utility value with a small effect ($d$ = 0.40) during the intervention whereas we can’t indicate any significant changes for intrinsic value and costs. The students’ trait values develop negatively. Thus, hypothesis 1 could not be confirmed. We assumed that the state value of each scale mediates the development between the corresponding trait value. We used bootstrap confidence intervals (based on 5,000 bootstrap samples) to test interference for indirect effects. An indirect effect is significant if zero lies not within the 95% bootstrap confidence interval.

| | T1 | T2 | $t$ (85) | Cohens' $d$ |
|---|---|---|---|---|
| AV | 4.17 (1.20) | 3.72 (1.34) | -4.89* | 0.53 |
| IV | 3.39 (1.42) | 3.19 (1.36) | -2.35 | 0.25 |
| UV | 4.32 (1.10) | 3.94 (1.11) | -3.68* | 0.40 |
| CO | 3.18 (1.19) | 3.31 (1.16) | 1.78 | 0.19 |

Table 2: Mean scores (standard deviations) and t-test results of the four task value components at the two measure points for trait values, $N$=86, answers between *1=not at all true* and *6=totally true*, $*p < .001$

For all models, we could find highly significant standardized total effects and direct effects (see table 3). The total effect describes how the trait values at T2 are affected by the trait values at T1 if the mediator is not statistically controlled. Answering the second research question by analyzing the total effects, the data unveil that students' trait values after the intervention are significantly and positively affected by their prior trait values for all value components. Thus, hypothesis 2 could be confirmed.

| | total | direct | indirect | | Path a | Path b |
|---|---|---|---|---|---|---|
| | β | β | β | 95% CI | β | β |
| AV | 0.77*** | 0.71*** | 0.07 | [-0.007, 0.158] | 0.40*** | 0.17 |
| IV | 0.84*** | 0.79*** | 0.05* | [0.014, 0.116] | 0.32** | 0.16* |
| UV | 0.62*** | 0.56*** | 0.06* | [0.017, 0.129] | 0.26** | 0.23** |
| CO | 0.84*** | 0.84*** | -0.01 | [-0.061, 0.047] | 0.37*** | -0.03 |

Table 3: standardized total, direct and indirect effects, and path coefficients of the mediation for all four models and their significance, $N$=94, $*p < .05$, $**p < .01$, $***p < .001$, Path a = Prediction T1 → S, Path b = Prediction S → T2

To answer the third research question, we focus on the paths from T1 to S (path a) and from S to T2 (path b) and the direct and indirect effects of each model (see figure 2). We find the path from T1 to S significant for all models, which refers to a positive top-down causality for all value components. Thus, we could confirm hypothesis 3.a. The path from S to T2 is only for intrinsic value and utility value significant, which reveals for these both components a positive bottom-up causality. Thus, hypothesis 3.b could partly be confirmed. The indirect effect is the product of path a and path b. The analysis reveals a significant indirect effect for the mediation of intrinsic value and utility value. The direct effect describes the effect between the trait values from T1 to T2 while controlling the mediator S. Because the direct effects are still significant the state value for these both components does not fully mediate the relationship between the trait values before the students worked on modelling tasks with experiments and their trait values afterwards. Therefore, hypothesis 3.c could be confirmed partly.

## DISCUSSION

Regarding the development of the students' trait values, we compared component-wise the means of the two measurement points T1 and T2. Against our hypothesis and findings of other studies (e.g., Beumann, 2016) the students' trait values decreased and costs increased after working on modelling tasks with experiments. So, it is questionable whether modelling tasks with experiments have the anticipated motivating effect. However, it must be noted that without a control group this statement is still weak. As the utility value decreases strongest, one could think that students do not consider modelling tasks with experiments as useful, as these tasks are not seen as typical tasks for mathematics lessons and therefore not useful. Also, a fatigue effect must be considered, as there are three consecutive mathematics lessons with the same structure. Our results rather support the findings of Krawitz and Schukajlow (2018) that students do not value modelling tasks compared to other tasks. Regarding the second research question we find a positive effect: the trait values after the intervention were positively affected by the trait values at the beginning. That supports the assumption of trait values to be rather stable personal features. The third research question investigates the role of the state value of each component as a mediator for the relationship between the students' trait value before working on the modelling tasks with experiment and their trait value afterwards. The top-down causality could be identified for all four components as the path from trait values in the beginning on state values was significant. The bottom-down causality could only be identified for intrinsic value and utility value, meaning that state value influenced the trait value in the end solely for these components of the task value. We were able to confirm the assumption of a mediation for intrinsic value and utility value. Given that the bottom-up and top-down causalities proposed by Moeller et al. (2022) is relatively recent, research is still in its nascent stages. Therefore, the strength respective the context independence of the trait-state relationship as top-down and bottom-up in the context of mathematics can be questioned. In the realm of mathematics in general students' motivation is commonly regarded stable (Lazarides et al., 2019).

The current study is limited in its statements as the size of the sample group is small considering the analyzation method. Moreover, modelling with experiments is a very specific context, which might influence the results and future studies should use a more innocuous area of mathematics. In general, the current study has shown that the framework from Moeller et al. (2022) is suitable to explain developments of students' values in mathematics. Therefore, the current study contributed to the field in terms of confirming the fitting of the trait and state relation regarding subjective task values in mathematics. Our ongoing research will focus on possible relations between the development of trait values, state values and students' performance in modelling tasks.

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

## References

Beumann, S. (2016). Versuch´s doch mal: Eine empirische Untersuchung zur Förderung von Motivation und Interesse durch mathematische Schülerexperimente. Dissertation. Ruhr-University Bochum.

Blum, W., & Leiß, D. (2007). How do Students and Teachers Deal with Modelling Problems? In C. Haines, P. Galbraith, W. Blum & S. Kahn (Eds.) *Mathematical Modelling: Education, Engineering, and Economics* (pp. 222–231). Elsevier.

Dietrich, J., Moeller, J., Guo, J., Viljaranta, J., & Kracke, B. (2019). In-the-Moment Profiles of Expectancies, Task Values, and Costs. Frontiers in Psychology, 10, 1662.

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

Eccles, J. S., & Wigfield, A. (2023). Expectancy-value theory to situated expectancy-value theory: Reflections on the legacy of 40+ years of working together. Motivation Science, 9(1), 1–12.

Ganter, S. (2013). Experimentieren - ein Weg zum Funktionalen Denken: Empirische Untersuchung zur Wirkung von Schülerexperimenten. Kovač.

Gaspard, H., Dicke, A.-L., Flunger, B., Schreier, B., Häfner, I., Trautwein, U., & Nagengast, B. (2015). More value through greater differentiation: Gender differences in value beliefs about math. Journal of Educational Psychology, 107(3), 663–677.

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

Krawitz, J., & Schukajlow, S. (2018). Do students value modelling problems, and are they confident they can solve such problems? Value and self-efficacy for modelling, word, and intra-mathematical problems. ZDM, 50(1–2), 143–157.

Lazarides, R., Dietrich, J., & Taskinen, P. H. (2019). Stability and change in students' motivational profiles in mathematics classrooms: The role of perceived teaching. Teaching and Teacher Education, 79, 164–175.

Ludwig, M. & Oldenburg, R. (2007) Lernen durch Experimentieren. Handlungsorientierte Zugänge zur Mathematik. mathematik lehren 141, 4–11.

Moeller, J., Viljaranta, J., Tolvanen, A., Kracke, B., & Dietrich, J. (2022). Introducing the DYNAMICS framework of moment-to-moment development in achievement motivation. Learning and Instruction, 81, 101653.

Niss, M. (1994). Mathematics in society. In R. Bieler, R.W. Scholz, R. Sträßer & B. Winkelmann (Eds.) Didactics of Mathematics as a Scientific Discipline (pp. 367-378). Kluwer Academic Publishers.

OECD (2017). PISA 2015 assessment and analytical framework. OECD Publishing.

Wigfield, A., & Eccles, J. S. (2000). Expectancy–Value Theory of Achievement Motivation. Contemporary Educational Psychology, 25(1), 68–81.

# MATHEMATICAL KNOWLEDGE FOR TEACHING COLLEGE ALGEBRA AT COMMUNITY COLLEGES

Vilma Mesa[1], Inah Ko[1], Irene Duranczyk[2], Patrick Kimani[3], Laura Watkins[3], Dexter Lim[2], Mary Beisiegel[4], April Ström[3], Bismark Akoto[2]

[1]University of Michigan, [2]University of Minnesota, [3]Maricopa Community Colleges, [4]Oregon State University

*We report on the relationship between community college instructors' performance on an instrument measuring mathematical knowledge for teaching college algebra with their teaching experience and their frequency of using specific tasks of teaching. The findings support the argument that the instrument assesses critical knowledge for teaching college algebra. We propose further work based on these findings.*

Studying the relationship between teacher knowledge and what teachers do in their classroom has a long history in mathematics education. Successful efforts to establish this relationship range from small-scale qualitative approaches in multiple settings to large-scale investigations, mainly in K-8 education, which developed instruments to measure both teachers' knowledge and their work in the classroom. One of the major limitations to replicate such findings in other contexts is the lack of reliable instruments that can measure such knowledge, especially in community college settings. Consequently, our project developed an instrument to assess mathematical knowledge for teaching college algebra among instructors who teach at community colleges in the United States. Community colleges are non-university tertiary institutions that enrol 38% of all undergraduate students in the country (American Association of Community Colleges, 2023) and offer opportunities for remediation, vocational training, worker retraining, general education, and transfer to four-year undergraduate institutions (Mesa, 2017). Our main goal is to establish a measure that can be used to determine the impact of professional development programs that target mathematical knowledge for teaching. In this report, we briefly describe the instrument we developed along with information that supports the argument that the instrument assesses mathematical knowledge for teaching college algebra in this context.

## RELEVANT LITERATURE

Most research seeks to understand the nature and composition of mathematical knowledge for teaching (MKT). Ball and colleagues (2008) proposed a framework for the MKT construct that accounted for multiple components of that knowledge (e.g., of students, of curriculum, and of mathematical horizon). Their work identified a relationship between teachers' knowledge about mathematics and its teaching, and the quality of their work in the classroom and that the quality of instruction had a positive impact on student learning in K-8 education (Hill et al., 2008). However, attempts to make connections between teachers' mathematical knowledge and their teaching experience measured as years in the profession have yielded mixed results. Hill (2010)



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 233-240). PME.

found no correlation between elementary teachers' overall years of mathematics teaching and their MKT scores. Krauss et al. (2008) found no relationship between years of experience teaching mathematics and performance on pedagogical content knowledge (PCK) and content knowledge (CK) assessments for secondary teachers, but those teachers prepared in pedagogy and mathematics content, when controlling for CK, demonstrated better PCK performance. Their study showed that beliefs rather than years of experience played a larger role in PCK performance. Some studies have found a positive relationship between MKT scores and experience teaching specific courses. For example, Herbst and Kosko (2014) found that high-school teachers who had taught geometry for at least three years had better scores on an MKT instrument designed to measure MKT in geometry compared to those instructors who had not taught such course. A data analysis from high-school algebra teachers who responded to an MKT instrument that assessed knowledge for teaching high-school algebra showed that respondents who had taught courses beyond algebra (e.g., calculus) performed better in the instrument than teachers who had not (Ko et al., 2021). Similar connections were also found by Hill (2007): middle school teachers with more experience teaching higher grades had higher MKT scores.

## THEORETICAL UNDERPINNINGS

We assume that the knowledge needed for teaching is better assessed in practice and that it is closely linked to the content that students need to learn. In the community college context, presenting concepts to students via examples is a common instructional practice which demands that instructors engage in at least two distinct tasks of teaching: (1) choosing problems that exemplify mathematical notions and (2) understanding students' work to ascertain whether students have understood the material (Mesa & Herbst, 2011). Once instructors have students' utterances or written work, they engage in a new process that demands selecting a problem that would help students to clear misconceptions or that may create a specific dissonance in their knowledge. Following the theory of didactical situations (Brousseau, 1997), and in particular the work of Herbst and colleagues (e.g., Herbst & Chazan, 2012), we assume that in any instructional situation, instructors must manage the interactions between students and content, and have the dual responsibility of, on one hand, offering students work that will directly relate to learning a piece of mathematics, and on the other, analysing the way in which mathematics is being addressed in students' utterances or in written work produced while learning mathematics. Attending to practices related to these two tasks of teaching (*choosing problems* and *understanding students' work*) can shed light on the connections within mathematical knowledge for teaching.

Since instructors at community colleges typically use examples to anchor the presentation of the material and solve them collaboratively with students, we developed an instrument that assesses MKT in the context of college algebra at community colleges and hypothesized that the knowledge needed to engage in these two tasks is different (Mesa et al., 2023). We also assumed that instructors who would more frequently engage in these two tasks of teaching would have a higher MKT.

Because teaching is a practice, it is also worthwhile to know how teaching experience relates to performance in MKT. As the studies in the literature review section show, there does not seem to be an association between number of years teaching mathematics in general, but with experience gained through teaching advanced courses that can build specific tasks of teaching. Given this background, we investigated the following two questions: What is the relationship between community college instructor performance in the Mathematical Knowledge for Teaching Community College Algebra (MKT-CCA) test and (1) their teaching experience and (2) their reported frequency of use of activities related to the tasks of choosing problems and understanding student work?

## METHODS

We recruited college algebra instructors from two-year degree granting institutions in the United States by inspecting community college websites and then sending direct invitations to instructors for participation in this project. Respondents included faculty from 260 different community colleges (~22% of institutions) in 42 states with 50% enrolling a majority of non-White students. Forty-eight percent of the participants identified as male and 46% as female; in terms of race, 76% identified as White, 10% as Asian, 4% as Black, 2% as mixed, and 4% chose Other. Seventy-eight percent of the participants said they held full-time positions; 9% were on tenure track. The average number of years of teaching experience was close to 17 years (mean = 16.76, SD = 9 years; range: 1.5 to 47 years). The majority (63%) held a master's degree in mathematics, mathematics education or another mathematics-related field, and 12% held PhDs (about 5% were in Mathematics Education).

### Instruments

The MKT-CCA test consists of 55 items that aim to measure MKT across six hypothesized dimensions (Duranczyk et al., 2023; Mesa et al., 2023). An analysis of the psychometric properties of items using the Two-Parameter Logistic (2PL) Item Response Theory (IRT) recommended the removal of 17 items that had a low discrimination estimate of less than 0.61, resulting in 38 items (Mesa et al., 2024). An exploratory factor analysis with the remaining 38 items was conducted to examine the structure of the item responses. This analysis identified six items suggested to be loaded on different factors than those where all other items were loaded. We excluded them from subsequent analyses because they did not appear to add anything new to the intended MKT construct. We established a unidimensional construct (MKT) with the remaining 32 items, using the weighted least square mean and variance adjusted (WLSMV) estimator. The model suggested a good fit (RMSEA = 0.016, CFI = 0.981, TLI = 0.980) according to the thresholds considered for a good fit (RMSEA < 0.06, CFI > 0.95, TLI > 0.95, Hu & Bentler, 1999). Standardized items loadings were all significant and greater than 0.30, indicating that all the items significantly contribute to the common MKT construct.

The eight items assessing frequency of tasks of teaching use a 5-point Likert scale (1- *never*, *almost never, sometimes, often*, 5- *very often*) and are statements in the form of activities related either to choosing problems (e.g., Modifying problems from a textbook or from colleagues) or to understanding student work (e.g., Noticing that a student's mathematical approach to a problem is valid even though it is not standard in a college algebra course). An exploratory factor analysis with the eight items revealed a two-factor model. Three items reflecting the task of choosing problems loaded on one factor (C1: Modifying problems from a textbook or from colleagues, C2: Evaluating how well a problem meets your instructional goal, C3: Assisting a student in a small way, such as by giving a mathematical hint or asking a question, without solving the problem for them) and three items reflecting the task of understanding student work loaded on another factor (U1: Trying to understand how students came up with their answer to a problem, U2: Reading students' work to figure out their thinking process, U3: Noticing that a student's mathematical approach to a problem is valid even though it is not standard in a college algebra course). The remaining two items had loadings of less than 0.3 on either factor. After examining these items, we excluded them from further analysis as they deviated from either the task of *understanding student work* or *choosing problems*.

Participants' teaching experience was assessed using two items, one that asked the number of years of full-time-equivalent teaching experience in mathematics and another that asked for the number of times the participants had taught post-college algebra courses such as Calculus 1, 2, 3, Business Calculus, Differential Equations, Linear Algebra, etc. This item was assessed on a 4-point scale (0 - *Never taught*, *Taught less than 5 times*, *Taught at least 5 times but less than 10*, 3 - *Taught 10 or more times*).

**Analysis**

To answer the first research question, we predicted the MKT construct by the two variables of teaching experience. To answer the second research question, we predicted the MKT score by the frequencies of performing tasks that require understanding student work or choosing problems. All analyses were conducted in Mplus (Muthén & Muthén, 1998-2023). Latent constructs were identified by setting factor means to 0 and factor variances to 1.

## RESULTS

We began conducting a regression analysis predicting the MKT from the frequency of two tasks of teaching. In the model, one of the items (task_C3 in Figure 1) asking the frequency of choosing problems was re-coded to prevent zero-frequency cells when using the WLSMV estimator. The model fit the data well (RMSEA = 0.018, CFI = 0.978, TLI = 0.977) and it showed a significant effect of the task frequency of choosing problems on MKT, but not with understanding student work. This significant effect indicates that instructors who frequently do the task of choosing problems tend to have higher MKT-CCA scores. Specifically, a one unit increase in the factor of task frequency is associated with 0.4 increase in MKT-CCA score (mean = 0, SD = 1) ($B = 0.383$, SE = 0.134, $p = 0.004$).

Second, we incorporated two additional predictors—the number of years teaching mathematics and the number of times instructors taught courses beyond college algebra—into the previous model. In this model, MKT-CCA is predicted by the two tasks of teaching frequency factors, the number of years teaching mathematics, and the number of times teaching courses post-college algebra. Additionally, the two task frequency factors were set to be predicted by the number of times teaching courses post-college algebra. The estimated structural model diagram is presented in Figure 1.

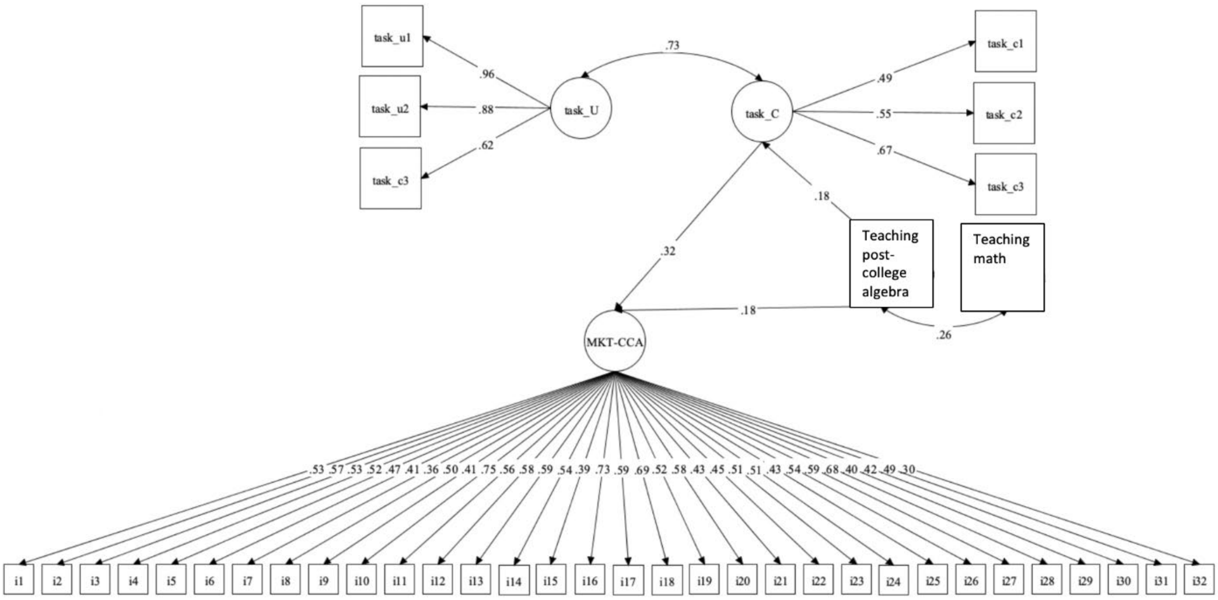



Figure 1. Structural model diagram showing only significant effects (standardized; $p < .05$)

The model fits the data well (RMSEA = 0.018, CFI = 0.974, TLI = 0.972). The result suggested a significant association between MKT-CCA and both the frequency of choosing problems ($B = 0.323$, SE = 0.138, p = 0.019) and the number of times teaching post-college algebra courses ($B = 0.178$, SE = 0.055, $p = 0.001$). When controlling for the number of times teaching post-college algebra courses, the effect of the task of choosing problems decreased in size and significance yet remained significant. This could be due to the relationship between the frequency of doing the task of choosing problems and the frequency of teaching post-college algebra courses, which is consistent with the significant effect of the number of times teaching post-college algebra courses on the frequency of doing the task of choosing problems ($B = 0.179$, SE = 0.066, p = 0.007). In contrast to the significant effect of the number of times teaching post-college algebra courses, there was no significant association between the MKT-CCA and the number of years teaching mathematics. This suggests that our MKT-CCA instrument measures a construct related to instructors' MKT associated with experience teaching post-college algebra courses.

## DISCUSSION

Our findings regarding the relationship between MKT-CCA and teaching experience, confirm prior findings (Ko et al., 2021) about the connection between MKT scores and teaching courses beyond college algebra. In that study, we found that with an instrument designed to assess knowledge for teaching 9$^{th}$ grade Algebra 1, community college instructors who had taught advanced courses scored better than those who had taught only college algebra. We believe that this finding speaks to a connection between the items assessed in this instrument and the knowledge that is acquired while teaching more advanced mathematics courses; such experiences allow teachers to reflect on the foundations that students will need to understand about mathematics that they will encounter later (the mathematics in the horizon, Ball et al., 2008). If instructors notice calculus students struggling with difference quotients, for example, they may realize the importance of bringing clarity to the notion of slope. We found that years of full-time experience teaching mathematics in general is not associated with MKT-CCA scores; this departs from prior finding, such as Hill (2010), who found that "more experienced [elementary] teachers—and particularly those with over 20 years of experience—have more MKT, and this overall relationship looks approximately linear" (p. 533). The discrepancy may rely on the nature of mathematics taught in elementary versus college courses.

For the second question we found a significant and positive relationship between the reported frequency of engaging in choosing problems and MKT-CCA performance. This suggests that the MKT construct that is being assessed by the instrument, might be associated with a task of teaching. Instructors with higher MKT might be more comfortable exercising choice for problems that might be more suitable to meet instructional goals, including those that address specific students' misunderstandings. A higher MKT-CCA score allows for identifying instructors who have a more nuanced set of instructional goals that will demand more attention when choosing problems. It may be that the instrument as designed might be capturing the knowledge needed to choose problems and thus, instructors who engage in that task more frequently will answer more items correctly. The model also showed that the frequency of use of the two tasks of teaching, *choosing problems* and *understanding student work*, are highly correlated. We believe that this is because as instructors spend time interpreting students' work, they might be using the opportunities to think about possible features of tasks that either led to student responses or think about tasks that could help students answer differently next time. But the reverse could also be true; as instructors choose problems, they might be generating hypotheses of the ways which students will answer them; once they receive responses, they will need to engage in understanding the work to decide whether their hypotheses were correct. Future studies will need to examine the mediating role of frequency of choosing problems and teaching courses post-algebra in relation to the MKT-CCA score, which was not possible to assess in this study because of the size of the analytical sample. Finally, we also found that the frequency of understanding student work did not predict the score in the MKT-CCA.

We think that this might be explained by the opportunities that community college instructors may have to interpret student work. As the number of students in these courses increase, colleges have been relying more on automated grading done by textbook companies (e.g., Pearson) that offer placement and homework systems (e.g., ALEKS) that eliminate the task of collecting and grading homework. It might be that instructors routinely have to understand and evaluate student work, making this practice similarly frequent across different teachers, which in turn does not significantly impact MKT. On the other hand, choosing problems, which is known to be less frequent, might potentially influence MKT due to the differing levels of experience among teachers. Alternatively, it could be possible that our instrument is less robust in assessing MKT used in understanding student work in the absence of choosing problems.

## CONCLUSION/IMPLICATIONS

This paper presents some preliminary results of an instrument designed to measure mathematical knowledge for teaching college algebra at the community college. The MKT-CCA links instructors teaching practices, particularly the reported frequency of choosing problems, with higher performance on the test. The MKT-CCA score also seems to increase based on the number of times instructors teach courses that build on college algebra skills. These findings, and further analyses, show promise for identifying instructor experiences that can be explored in professional development settings and help raise awareness of teaching practices that enhance mathematical knowledge for teaching.

### Acknowledgments

Funding provided by the National Science Foundation EHR Core Awards #2000602 (Michigan), #2000644 (Maricopa), #2000527 (Minnesota), and #2000566 (Oregon). Opinions, findings, and conclusions or recommendations expressed do not necessarily reflect the views of the Foundation.

### References

American Association of Community Colleges. (2023). *Fast Facts.* https://www.aacc.nche.edu/wp-content/uploads/2023/03/AACC2023_FastFacts.pdf

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

Brousseau, G. (1997). *Theory of didactic situations in mathematics* (N. Balacheff, Trans.). Kluwer.

Duranczyk, I., Mesa, V., Ko, I., & VMQI–AI@CC 2.0 Team. (2023). Mathematical Knowledge for Teaching for college algebra at community colleges. In M. Ayalon, B. Koichu, R. Leikin, L. Rubel, & M. Tabach (Eds.), *Proceedings of the 46th Conference of the International Gorup for the Psychology of Mathematics Education* (Vol. 2, pp. 251–258). University of Haifa, Israel: PME.

Herbst, P., & Chazan, D. (2012). On the instructional triangle and the sources of justifications for the actions of the mathematics teacher. *ZDM Mathematics Education*, *44*, 601-612. https://doi.org/10.1007/s11858-012-0438-6

Herbst, P.G., & Kosko, K.W. (2014). Mathematical knowledge for teaching and its specificity to high school geometry instruction. In J.-J. Lo, K. R. Leatham. & L. R. Van Zoest (Eds.), *Research trends in mathematics teacher education* (pp. 23–45).

Hill, H. (2007). Mathematical knowledge for middle school teachers. Implications for the No Child Left Behind policy initiative. *Educational Evaluation and Policy Analysis, 29*(2), 95–114.

Hill, H. C. (2010). The nature and predictors of elementary teachers' mathematical knowledge for teaching. *Journal for Research in Mathematics Education, 41*(5), 513–545.

Hill, H. C., Ball, D. L., & Schilling, S. G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education*, *39*(4), 372–400.

Hill, H. C., Blunk, M., Charalambous, C., Lewis, J., Phelps, G., Sleep, L., & Ball, D. L. (2008). Mathematical Knowledge for Teaching and the Mathematical Quality of Instruction: An exploratory study. *Cognition and Instruction*, *26*(4), 430–511.

Ko, I., Mesa, V., Duranczyk, I., Herbst, P., Kohli, N., Ström, A., & Watkins, L. (2021). Understanding the characteristics of mathematical content knowledge for teaching algebra in high schools and community colleges. *International Journal of Mathematical Education in Science and Technology.* https://doi.org/10.1080/0020739X.2021.2006348

Krauss, S., Baumert, J., & Blum, W. (2008). Secondary mathematics teachers' pedagogical content knowledge and content knowledge: Validation of the COACTIV constructs. *ZDM – Mathematics Education, 40*, 873–892.

Mesa, V. (2017). Mathematics education at public two-year colleges. In J. Cai (Ed.), *First compendium for research in mathematics education* (pp. 949-967). National Council of Teachers of Mathematics.

Mesa, V., Duranczyk, I., Wingert, K., & VMQI-AI@CC 2.0 Team. (2024). *Measuring mathematical knowledge for teaching college algebra at community colleges.* Research in Undergraduate Mathematics Education Conference, Omaha, NE, United States.

Mesa, V., & Herbst, P. (2011). Designing representations of trigonometry instruction to study the rationality of community college teaching. *ZDM The International Journal on Mathematics Education*, *43*, 41-52. https://doi.org/10.1007/s11858-010-0300-7

Mesa, V., Ko, I., Boeck, C., Duranczyk, I., Akoto, B., Lim, D., Kimani, P., Watkins, L., Ström, A., & Beisiegel, M. (2023). *A blueprint for measuring mathematical knowledge for teaching college algebra at community colleges.* 13th Congress of the European Society for Research in Mathematics Education, Budapest, Hungary.

Muthén, L. K., & Muthen, B. (1998-2015). *Mplus user's guide: Statistical analysis with latent variables, user's guide*. Muthén & Muthén.

# OPERATIONALIZING RE-PRESENTATION TO INVESTIGATE AND SUPPORT STUDENTS' COVARIATIONAL REASONING

Kevin C. Moore[1], Erin Wood[1], Sohei Yasuda[1], Irma E. Stevens[2], Biyao Liang[3], Halil I. Tasova[4]

[1]University of Georgia, [2]U. of Rhode Island, [3]The U. of Hong Kong, [4]CSU-San Bernardino

*Within the body of work on students' covariational reasoning, researchers have called for more explicit attention to the ways theoretical constructs are operationalized to develop characterizations of student thinking. Addressing this need, we outline how von Glasersfeld's (1991) notion of re-presentation—the act of reconstructing something previously experienced in its absence—has informed our research program on students' covariational reasoning. Specifically, we illustrate its multimodal use in framing claims regarding the extent a student has constructed a particular covariational relationship.*

## INTRODUCTION

Covariational reasoning refers to the mental operations involved in coordinating two quantities' magnitudes or values as they vary in tandem (Thompson & Carlson, 2017). Students' covariational reasoning remains a growing area of study due to researchers having illustrated its critical foundation for students constructing major algebra, function, calculus, and STEM concepts (Thompson & Carlson, 2017). Accordingly, researchers have provided a variety of models of student thinking, with each model entailing the use of theoretical constructs to make aspects of student thinking salient. For instance, Carlson et al. (2002) specified several mental actions associated with students' covariational reasoning. Similarly, Ellis et al. (2020) and Johnson (2015) have each characterized nuances in the ways students reason about covariation.

A by-product of growth in an area of study is that guiding theories and constructs become more or less noticeable as researchers develop more nuanced or detailed characterizations. For example, as researchers have developed more specified descriptions of the mental actions involved in students' covariational reasoning, macro-level constructs that focus on general properties or forms of reasoning have moved to the background. This progression is natural and often necessary, yet it has notable consequences (Tyburski et al., 2021). For one, it leaves unclear the ways in which macro-level constructs emerged and continue to inform research design or analysis. For another, it inhibits other researchers adopting the work for their own purposes. In a call to fellow researchers, Tyburski et al. (2021) argued these consequences negatively impact the accessibility of research to novice or outsider researchers.

We respond to this call by identifying the ways von Glasersfeld's (1991) notion of re-presentation—the act of reconstructing something previously experienced in its absence—has informed our research program on students' covariational reasoning. In



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 241-248). PME.

what follows, we first provide background information on re-presentation and students' covariational reasoning. We then discuss the explicit ways in which re-presentation has emerged and informed our research. Namely, we have used re-presentation to consider and frame the viability of our claims regarding students' covariational reasoning.

## RE-PRESENTATION

von Glasersfeld's notion of re-presentation emerged during his study of Piaget's genetic epistemology and as a distinction relevant to object permanence (von Glasersfeld, 1991, 1995). Re-presentation refers to the ability of an individual to construct a visualized image of an object in the absence of the relevant sensory material. von Glasersfeld emphasized the hyphenated form of re-presentation for two primary reasons. As the first reason, the hyphenated form reflects that to both von Glasersfeld and Piaget, re-presentation is an *active* attempt to present again. Because re-presentation involves regenerating a past experience or concept in the absence of the relevant figurative material, it is subject to and defined by the ways of operating available to the individual at that moment. Re-presentation does not produce a copy of the previous experience or concept, nor is it a simple recall of the previous experience as with a ready-made picture. Relatedly, because re-presentation is an active process, a researcher should not presume that the operations involved in re-presentation are equivalent to those used during the initial experience. This is particularly true when a large duration of time separates the two. As the second reason, von Glasersfeld's insistence on using the hyphenated re-presentation reflects his linguistics background. He desired to distinguish between *re-presentation* and *representation*. Whereas the former is a constructive process involving the enactment of conceptual structures, he defined the latter as something acting as a copy, a pointer, or something that stands in for something else (von Glasersfeld, 1995). For instance, one might say a displayed Cartesian line and the inscription "$y = 3x$" represent (without hyphen) a linear relationship, whereas a re-presentation (with hyphen) of a linear relationship involves enacting conceptual operations associated with the conceived relationship to regenerate associated figurative material. We expand on this example in the next section.

Further emphasizing its importance for the construction of concepts, von Glasersfeld described re-presentation as one of the key drivers of abstraction and learning. He considered the re-presentation of objects and conceptual structures to enable the construction of hypothetical situations not available on an experiential or sensorimotor basis. In his words, re-presentation enables thought experiments, and through affording processes of abstraction "thought experiments constitute what is perhaps the most powerful learning procedure in the cognitive domain" (von Glasersfeld, 1995, p. 69). As an apropos example, Steffe and colleagues' (Steffe & Olive, 2010) extensive research program on fractional reasoning illustrates that acts of re-presentation are inseparable from the construction of number and multiplicative reasoning.

## MAGNITUDES, OPERATIONS, AND COVARIATION

Research on covariational reasoning, or "reasoning about values of two or more quantities varying simultaneously" (Thompson & Carlson, 2017, p. 423), has primarily occurred within Thompson's quantitative reasoning paradigm. Informed by von Glasersfeld's radical constructivism and Piaget's genetic epistemology, Thompson defined a *quantity* as a measurable attribute of some situation (Thompson, 1989). Reflecting the theory's epistemological underpinning, Thompson emphasized that quantities and their relationships are cognitive constructions and thus idiosyncratic to the knower. Researchers have since adopted this perspective to develop insights into the quantities and covariational relationships students and teachers construct (see Thompson & Carlson, 2017 for a summary of this work). We focus on two aspects from this work in order to connect re-presentation to students' covariational reasoning.

Firstly, a fundamental distinction in Thompson's theory is that between quantitative operations and arithmetic operations (Thompson, 1989). The former refers to the mental operations involved in constructing a quantity and associated amountness, while the latter refers to numerical operations that define or calculate a quantity's measure or value. To clarify, consider using the inscriptions "2" or "6-4" to represent a measure or comparison between measures. Here, represent (no hyphen) is used in the sense of their standing in for or pointing to anticipated conceptual (quantitative) structures. Because re-presentation stresses the *enactment* of mental operations in order to reconstruct a conceptual structure (von Glasersfeld, 1995), re-presenting "2" involves reconstructing quantitative operations including creating and iterating a unit magnitude in the context of figurative material that permits those operations (e.g., a segment). With respect to the inscription "6-4", an act of re-presentation involves reconstructing those same operations for "6" and "4", and then reconstructing the operations involved in disembedding and measuring the magnitude by which the "6" length exceeds the "4" length (Thompson, 1989). Underscoring the difference between re-presenting operations and representing, we suspect the reader immediately understands "2" as representing the result of evaluating "6-4" without having to enact in re-presentation the operations represented by "6-4" or the additive difference of "2".

Secondly, Carlson et al. (2002) provided a framework of mental actions that specify several quantitative operations involved in covariational reasoning. For the purposes of this paper, we draw attention to *direction of change* and *amount of change* operations. Direction of change involves conceiving variation in one quantity's magnitude in tandem with variation in another quantity's magnitude. For instance, in the context of counter-clockwise circular motion from a 3 o'clock position, the height above the circle's center increases as the arc length traversed increases (Figure 1). Here, the quantities' magnitudes are paired while each quantity's magnitude is compared across states via a gross comparison with its previous state. Amount of change involves further quantifying quantities' covariation by systematically comparing the accumulation of each quantity. As an example, one can capture the arc length's accumulation by constructing and iterating a unit arc length. Pairing height with the arc length's accumulation, the individual can construct and additively compare not only successive heights, but also the successive variations in height (Figure 1). Here, the variations in both quantities' magnitudes are coordinated, with one quantity's variation remaining equivalent in magnitude (i.e., equal, successive increases) while the variation in the other quantity's magnitude is compared across states via a gross comparison (i.e., the increase is decreasing). We underscore that this illustration centers quantitative operations, magnitudes, and associated figurative material, as opposed to specified values, inscriptions representing those values, or arithmetic operations involving values. Each are critical for mathematical development and communication, but acts of re-presentation involve the reconstruction of the former.

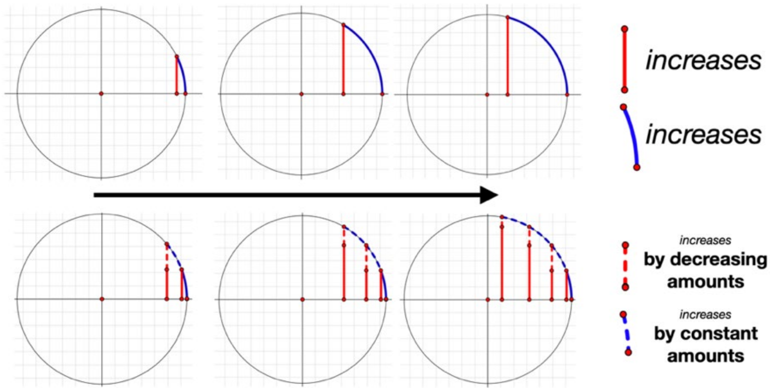



Figure 1: Direction of change (top) and amount of change (bottom).

## RESEARCH CONTEXTS

This paper emerged from the empirical work of building accounts of student thinking in the context of major algebra, pre-calculus, and calculus ideas. The primary attention of this work has been understanding, engendering, and supporting students' and teachers' quantitative and covariational reasoning. The work involved a series of teaching experiments with middle-grade, secondary, and undergraduate students and teachers. A *teaching experiment* is a qualitative design-based research methodology that involves constructing and testing hypothetical models of student thinking (Steffe & Thompson, 2000). Analytic methods of conceptual analysis (Steffe & Thompson, 2000) in combination with generative and convergent coding (Corbin & Strauss, 2008) accompanied the teaching experiments. It was during the iterative execution and analyses of the teaching experiments that re-presentation emerged as a useful construct, and we point the reader to Stevens (2019), Liang and Moore (2021), and Moore et al. (2022) for specified accounts of and references to this empirical work and findings.

## RE-PRESENTATION AND CLAIM VIABILITY

The initial need for re-presentation as an explanatory construct emerged when our research team noticed a similar phenomenon during a series of studies: a student had engaged in activity that strongly suggested their having constructed a stable understanding of some covariational relationship, but their actions during subsequent tasks suggested otherwise. For example, in exploring circular motion, we experienced students repeatedly producing diagrams consistent with Figure 1 along with the appropriate verbal descriptions. The fluidity of their actions led us to believe they had constructed a sophisticated and stable covariational relationship. However, the student would experience difficulties when prompted to construct a Cartesian graph of the relationship, or to choose two segments that match the covariational relationship from a collection of varying segments. The difficulties occurred in two primary ways.

In some cases, a student's difficulty would occur when they attempted to return to and regenerate the original situation and relationship in the presence of a new task. As an example, a student named Lilly attempted to regenerate the relationship illustrated in Figure 1 when attempting to determine which two segments from a collection of varying segments captured the sine relationship (Figure 2). Illustrated in detail in Liang and Moore (2021), Lilly desired to use the displayed circle to regenerate the relationship she previously determined as "sine" so she could compare it with how chosen segment-pairs covaried. However, she experienced difficulty regenerating the relationship unless the researchers provided figurative material (e.g., marks to visually denote amounts of change) to support her in making quantitative comparisons.

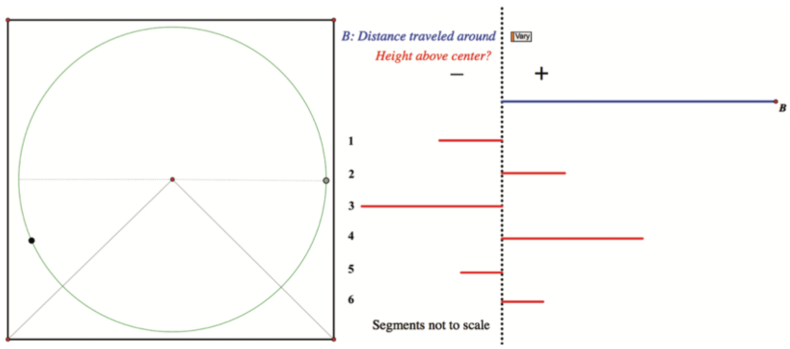



Figure 2: Choosing from six (red) varying segments (Liang & Moore, 2021, p. 300).

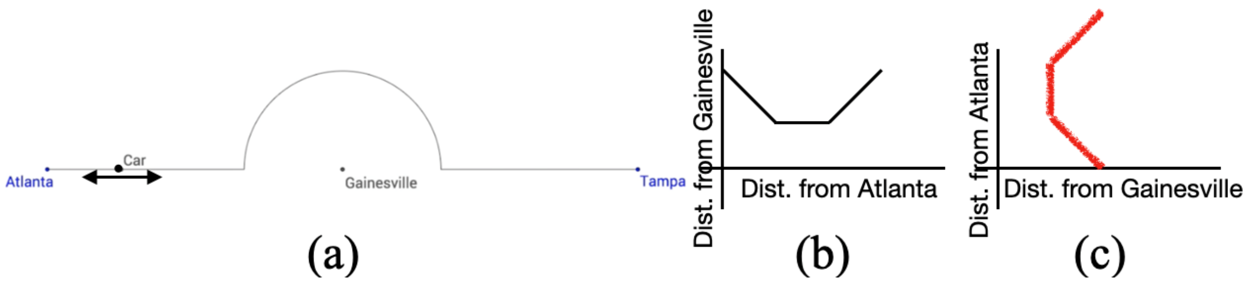



Figure 3: The (a) task situation and (b-c) normative graphs.

In other cases, a student would return to and regenerate the original situation and relationship without trouble, but the student would experience a difficulty regenerating a previously constructed relationship using the figurative material of a new task. As an example, after determining a covariational relationship in a situation and constructing a graph of that relationship by re-presenting the quantities' covariation (Figure 3a-b), Moore et al. (2019) reported on a student abandoning the construction of the graph in an alternative Cartesian coordinate orientation (i.e., the axes swapped, Figure 3c). The student, Patty, experienced no issues regenerating the covariational relationship in the situation or using the initial coordinate orientation, but she perceived creating a graph in the new coordinate orientation as requiring drawing it "right-to-left." She claimed such a graph is "backwards" and must be incorrect because of that feature.

The frequency of cases like these in tandem with the students' experienced difficulties being sustained and significant led us to question the extent we could claim the students' reasoning foregrounded covariational reasoning. In Lilly's case, we perceived her difficulties in re-presenting the relationship in her previously experienced context to be a contraindication of such reasoning. In Patty's case, her difficulties in re-presenting the relationship under a new coordination orientation were also a contraindication of such reasoning. We thus searched for a construct that could help us not only characterize each case, but also differentiate between them.

We do not recall the first instance in which we came across re-presentation as a potential tool. But, it became clear that re-presentation would be a useful tool when a research team member was in the depths of her dissertation work and needed to distinguish between students' uses of formulas as inscriptions capturing arithmetic rules between values or as symbolizing quantitative operations relevant to a dynamic geometric object (Stevens, 2019). Upon coming across re-presentation, our team returned to our data to engage in further rounds of conceptual analysis. In doing so, re-presentation's dual emphasis on the availability of figurative material and the reenactment of conceptual operations provided us a way to situate our claims regarding students' reasoning so that we considered them viable. Here, our use of viable is compatible with Steffe and Thompson (2000). We consider a claim viable if it is both an adequate hypothetical account of student thinking and it is specified enough to convey both affordances and constraints in their reasoning.

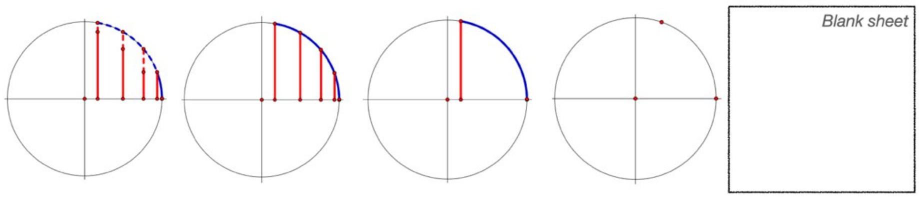



Figure 4: Varying the provided figurative material.

Reflecting on the cases above and considering the dual emphasis of re-presentation, we can explore indications and contraindications regarding students' covariational reasoning in two ways after a student has engaged in activity that we take as providing evidence of covariational reasoning. Firstly, as researchers, we can prompt a student to re-present their actions within the same context or phenomenon as previously experienced. Furthermore, we can vary the amount of figurative material provided to them. For instance, after a researcher has evidence a student has constructed the relationship consistent with Figure 1, during a subsequent task the researcher could prompt the student to reconstruct that relationship, and they could do so in a way that provides a range from a completed diagram to a blank sheet of paper (Figure 4). Returning to Lilly, when only provided a dynamic point on a circle, she could not re-present her previously constructed relationship. But, when provided the collection of heights all at once, she was able to re-present her previously constructed relationship.

Secondly, we can prompt students to re-present their actions within a different (or series of different) context(s) or phenomenon(s). For instance, after a researcher has evidence a student has constructed a relationship in a phenomenon (e.g., circular motion or a road trip), the researcher could prompt the student to reconstruct that relationship within a variety of Cartesian orientations (e.g., Figure 3), alternative coordinate systems (e.g., polar coordinates), or number line situations (e.g., Figure 2). A researcher can also vary the amount of figurative material available in the new contexts or phenomenon. For instance, in moving to a different coordinate system (e.g., polar coordinates), a researcher may or may not provide a quantity's variation partitioned (e.g., a marked grid). Such moves support a researcher in differentiating between a student's understanding of a particular covariational relationship and their generalized understanding of the coordinate system's quantitative structure. Returning to students like Patty, if a student considers drawing a graph "left-to-right" to be absolutely necessary, then no amount of figurative material would immediately support them in drawing and accepting a normative graph in the given orientation. On the other hand, in the original Cartesian orientation, Patty was able to re-present her relationship. Illustrating how re-presentation supports a researcher in situating their claims, Patty's actions indicate that she had constructed a covariational relationship she could re-present graphically, but her Cartesian graphing meanings entailed properties of movement that did not support her in doing so for a particular orientation.

## CLOSING COMMENTS

von Glasersfeld's notion of re-presentation enables a researcher to situate their claims regarding a student's covariational relationship with respect to 1) the amount of figurative material necessary to re-present the relationship, 2) their ability to re-present the relationship in other contexts and phenomenon, and 3) a combination of the two. By designing task environments sensitive to these re-presentational framings, we can systematically pursue indications and contraindications of students having constructed particular covariational relationships based on their capacity to re-present those relationships. Importantly, adopting a re-presentation framing has increased our sensitivity to the properties and features that students abstract from their their initial construction of a quantitative or covariational relationship. This supported sensitivity underscores von Glasersfeld's framing of re-presentation as a driver of learning.

On the topic of learning, it is important to note that students' re-presentational activity can and does change over time. What a student is able to re-present from one day to another might not be available to them at a later time. Likewise, what a student cannot re-present at one moment in time may become available to them in re-presentation at another moment in time. This phenomenon is inherent to the learning process and cognitive development (Steffe & Olive, 2010; von Glasersfeld, 1995), and it suggests that instruction and curricular materials should give direct attention to engendering and supporting cycles of students' re-presentational activity. Not only is re-presentation a driver of abstraction, it is a precursor to meaningful symbolization, and thus it provides a foundational springboard to an individual's mathematical development. By responding to the call by Tyburski et al. (2021), we hope to not only provide insights into how re-presentation has emerged in our research, but also invite conversation about how it might inform the teaching and learning of mathematics more broadly.

## References


Carlson, M. P., Jacobs, S., Coe, E., Larsen, S., & Hsu, E. (2002). Applying covariational reasoning while modeling dynamic events: A framework and a study. *Journal for Research in Mathematics Education*, *33*(5), 352-378.

Corbin, J. M., & Strauss, A. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (3rd ed.). Sage.

Ellis, A. B., Ely, R., Singleton, B., & Tasova, H. (2020). Scaling-continuous variation: supporting students' algebraic reasoning. *Educational Studies in Mathematics*.

Johnson, H. L. (2015). Secondary students' quantification of ratio and rate: A framework for reasoning about change in covarying quantities. *Mathematical Thinking & Learning*, *17*(1), 64-90.

Liang, B., & Moore, K. C. (2021). Figurative and operative partitioning activity: A student's meanings for amounts of change in covarying quantities. *Mathematical Thinking & Learning*, *23*(4), 291-317.

Moore, K. C., Liang, B., Stevens, I. E., Tasova, H. I., & Paoletti, T. (2022). Abstracted Quantitative Structures: Using Quantitative Reasoning to Define Concept Construction. In G. Karagöz Akar, İ. Ö. Zembat, S. Arslan, & P. W. Thompson (Eds.), *Quantitative Reasoning in Mathematics and Science Education* (pp. 35-69). Springer International.

Moore, K. C., Stevens, I., Paoletti, T., Hobson, N., & Liang, B. (2019). Pre-service teachers' figurative and operative graphing actions. *The Journal of Mathematical Behavior*, *56*.

Steffe, L. P., & Olive, J. (2010). *Children's Fractional Knowledge*. Springer.

Steffe, L. P., & Thompson, P. W. (2000). Teaching experiment methodology: Underlying principles and essential elements. In R. A. Lesh & A. E. Kelly (Eds.), *Handbook of research design in mathematics and science education* (pp. 267-307). Erlbaum.

Stevens, I. E. (2019). *Pre-service teachers' constructions of formulas through covariational reasoning with dynamic objects* [Ph.D. Dissertation]. University of Georgia: USA.

Thompson, P. W. (1989, March 27-31). *A cognitive model of quantity-based algebraic reasoning* Annual Meeting of the AERA, San Francisco, CA.

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.

Tyburski, B., Drimalla, J., Byerley, C., Boyce, S., Grabhorn, J., & Moore, K. C. (2021). From theory to methodology: Guidance for analyzing students' covariational reasoning. In D. Olanoff, K. Johnson, & S. M. Spitzer (Eds.), *Proceedings of the 43 annual meeting of the PME-NA*.

von Glasersfeld, E. (1991). Abstraction, re-presentation, and reflection: An interpretation of experience and of Piaget's approach. In L. P. Steffe (Ed.), *Epistemological foundations of mathematical experience* (pp. 45-67). Springer.

von Glasersfeld, E. (1995). *Radical constructivism: A way of knowing and learning*. Falmer Press.

# UNVEILING PROSPECTIVE TEACHERS' CONCERNS: USING A GUIDED REFLECTION PROCESS AS PART OF MATHEMATICS TEACHER EDUCATION

Galit Nagari-Haddif, Ronnie Karsenty, and Abraham Arcavi

Weizmann Institute of Science, Israel

*This report presents insights from a study focusing on guided reflection processes of prospective mathematics teachers. The study explores how two prospective teachers, in their practicum year, use the Six Lens Framework (SLF) - a tool originally designed for the professional development of practicing teachers. The findings highlight unique concerns of future teachers and demonstrate how reflection focused on distinct aspects of practice can contribute to their learning from observed lessons. We present various kinds of analyses preformed on the teachers' accumulating reflections and discuss the differences and similarities between the two teachers' implementation of SLF. Finally, we suggest implications for teacher education.*

## INTRODUCTION

Teacher reflection is a complex and multifaceted process consisting of describing and analyzing teaching events in order to develop awareness to currently-held practices and possibly reshape them (Finlay, 2008). By judiciously inspecting instructional strategies and classroom dynamics, teachers may enhance understanding of their own educational philosophies and methods (e.g., Bjuland et al., 2012) and also their perception of student needs. Reflection may result in more effective teaching approaches and better student outcomes and may lead to the refinement of teacher identity (e.g., Rhoads & Weber, 2016). Thus, reflection is increasingly becoming a focus of research on practicing mathematics teachers' professional development (e.g., Nurick at al., 2022).

Several resources can facilitate reflection, among them the Six-Lens Framework (SLF; details follow) which was successfully implemented with practicing teachers (Karsenty & Arcavi, 2017). In this paper, we report on prospective teachers (PSTs)' experiences with SLF-based reflection. The needs of prospective and practicing mathematics teachers differ in many respects: whereas PSTs seek foundational knowledge and experiences, practicing teachers pursue ongoing professional growth opportunities linked to their actual teaching realities (Lin & Rowland, 2016). PSTs do not yet consider themselves as "genuine" teachers, since they have not experienced the full realities of classrooms, except for their own encounters as pupils (e.g., Chamoso et. al, 2012). Eliciting and supporting PSTs' reflection, based on SLF, may yield insightful outcomes on the nature of PSTs' needs, concerns and identities (as indicated, for example, by some results already reported in Chikiwa & Graven, 2023).



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 249-256). PME.

## THE SIX-LENS FRAMEWORK

The SLF tool, aimed at enhancing reflection on multiple aspects of mathematics teaching practices, was developed within a professional development (PD) project called VIDEO-LM (Viewing, Investigating and Discussing Environments of Learning Mathematics; Karsenty & Arcavi, 2017). In the project's PD sessions, groups of mathematics teachers collaboratively analyze videos of authentic mathematics lessons taught by unknown colleagues. Studies around the project (e.g., Schwarts & Karsenty, 2020) show that the participating teachers delve into profound conversations about fundamental aspects of their profession. The SLF includes the following lenses, used for observing and discussing lessons (Karsenty & Arcavi, 2017):

1. Mathematical and meta-mathematical ideas related to the lesson's topic;
2. Explicit and implicit goals that may be ascribed to the teacher for the lesson;
3. The tasks selected by the teacher and their enactment in class;
4. The nature of the classroom interactions;
5. The teacher's dilemmas and decision-making processes;
6. The teacher's beliefs about mathematics, its learning and its teaching, as inferable from the teacher's actions.

The present study adapted the SLF tool for PSTs, who in addition to observing videotaped lessons of unknown teachers, observed live mathematics lessons in the same classes over an extended period of time, in the school where they performed their practicum. We focused on the following research question: What are the characteristics and foci of PSTs' reflection processes, as detected by their use of the SLF tool?

## METHODOLOGY

*Setting*. Data for this study were gathered during 2021-2 at a teacher education college, within the mathematics teaching track. The PSTs involved in the study were introduced to the SLF tool in their first year, in a methods course taught by the first author of this paper. In this course, they engaged in watching and analyzing videos of mathematics lessons from the VIDEO-LM website, after which they submitted written reflections. In the subsequent year, this group of PSTs entered their practicum, which took place in a local high school (grades 7-12). Typically, a day in the practicum program began with an opening meeting, followed by 4 hours of classroom observations in pairs or in small groups, guided by pre-assigned viewing lenses (usually different lenses each time), and concluding with a plenary for sharing insights. Each PST maintained a 'personal reflective blog', later submitted to the course instructor for feedback. A certain degree of flexibility was enabled, for example PSTs were not directed or limited in regards to the length of reflections under each lens, and were occasionally also free to replace the lenses they were assigned, in cases when they found it conducive.

*Participants*. The study subjects comprised all PSTs in the mathematics track cohort beginning in 2021, 11 in total (all females). All studied towards a teaching certificate in secondary mathematics, and were about 22 years old, on average. In the practicum, they were organized into four fixed subgroups, reshuffled in the second semester, and partnered with two supervising teachers for classroom observations. In this report we focus on two participants, Kim and Mia, to illustrate distinctly different cases of PSTs' reflection processes. Kim is a motivated student with high verbal skills, whereas Mia, a less verbal PST, began the practicum year with what seemed as low motivation. This was also reflected in their blogs, with Kim's being considerably longer than Mia's. Thus, they were selected as representing two 'extreme' sides of a spectrum.

*Data collection and analysis*. We used two data sources: (a) all reflections uploaded by the PSTs to the course website along the practicum year; and (b) two semi-structured 1-hour interviews, in the middle and in the end of the year, conducted with each PST by the first author. The interviews elicited the PST's perspectives on the practicum as a whole and on elements such as the use of SLF, group work, and more. All interviews were recorded and transcribed. For analyzing the written reflections, we first segmented the texts into content units we called 'statements'. Then, we reviewed a sample of statements, assigning each one two labels, one pertaining to the lens utilized by the PST, and the other, in a bottom-up process, pertaining to emerging issues (e.g., class atmosphere, explaining mathematical concepts). Next, we reached consensus on issue categories, resulting in a revised coding scheme, which was systematically reapplied to all statements generated by each PST. Finally, we performed two additional types of analysis: (I) separating statements into *generic vs. mathematical* issues; and (II) identifying statements featuring *a personal aspec*t, i.e., when the PST reflected on her own preferences, difficulties, etc., following the lesson she observed. We determined the distributions of statements according to each of the 4 types of analysis (lens, issue category, mathematical vs. generic, personal aspect). The total number of statements produced by Kim and Mia and analyzed in the study was 277 and 39, respectively. For the interviews, a content analysis was performed, scrutinizing the PST's experiences regarding the application of SLF and the overall process they had undergone. The interview analysis served to support and explain findings from the reflection analyses.

## FINDINGS

*Frequency of lenses employed by the two PSTs*. The distribution of the six lenses in the reflections of Kim and Mia is presented in Table 1. As can be seen, the lens most commonly employed was the lens of Interactions, whereas the lens of Mathematical ideas was least frequently used by both Mia and Kim. The excerpt below, from Kim's first interview, sheds some light on this phenomenon:

> The lens that is most difficult for me is the mathematical ideas. It's very hard in its analysis, beyond [the challenge of] connecting to it. The lens of goals is also hard for me because it's more general than the other lenses. I connect more to the other four lenses. They're different, but in the end they are very similar. They're pretty much based on the same principle. If I take the lenses of interactions, dilemmas, beliefs, and tasks, it's like I'm looking at the same thing but taking it to a different place.

| Lens | Kim | Mia |
|---|---|---|
| Interactions | 80 (28.9%) | 15 (38.5%) |
| Beliefs | 72 (26.0%) | 8 (20.5%) |
| Tasks | 43 (15.5%) | 0 (0%) |
| Teacher dilemmas and decision-making | 39 (14.1%) | 6 (15.4%) |
| Goals | 38 (13.7%) | 10 (25.6%) |
| Mathematical and meta-mathematical ideas | 5 (1.8%) | 0 (0%) |
| Total | 277 (100%) | 39 (100%) |

Table 1: The distribution of lenses in the two PSTs' reflections.

We suggest that the 'generic' lens of Interactions was easier and more appealing for the PSTs to use, for two reasons: (1) it demands less expertise in the disciplinary intricacies of mathematics teaching, an expertise often characteristic of experienced teachers but not yet available to PSTs; and (2) it allows PSTs to focus on issues that commonly trouble them, such as pedagogical strategies and class management. This conjecture is reinforced by looking at the content categories, as we present next.

*Frequency of topics raised in the PSTs' reflections*. The bottom-up coding process described above yielded 5 main categories (recurring in at least 4% of the statements) and 11 smaller categories (appearing in less than 3% of the statements). Due to space limitations, in this report we combine all the smaller categories under the category of "Other", but we note that some of them pertain to generic issues (e.g., students' confidence and self-image), while others are related to mathematics (e.g., students' mathematical mistakes). The distribution of the content categories in the reflections written by Kim and Mia is presented in Table 2.

| Categories | Kim | Mia |
|---|---|---|
| General pedagogical strategies | 98 (35.4%) | 13 (33.3%) |
| Mathematical teaching strategies | 90 (32.5%) | 9 (23.1%) |
| Classroom management | 21 (7.6%) | 6 (15.4%) |
| The nature of teacher-led discussions | 13 (4.7%) | 3 (7.7%) |
| Teacher caring and classroom atmosphere | 12 (4.3%) | 3 (7.7%) |
| Other issues | 43 (15.5%) | 5 (12.8%) |
| Total | 277 (100%) | 39 (100%) |

Table 2: The distribution of content categories in the two PSTs reflections.

As the table shows, the highest attention was given to general pedagogical strategies (about a third of the statements, for both Mia and Kim). Next is the category of mathematical teaching strategies, again a salient focus for both PSTs, and the third largest category is classroom management.

The analysis separating statements into mathematical vs. generic issues revealed that a majority of both PSTs' reflections pertained to generic issues (61% for Kim and 66.7% for Mia). This aligns with the finding regarding the PSTs favoring the use of the Interactions lens.

*Personal aspect*. We found that 16.6% of Kim’s statements were coded as featuring a personal perspective, while none of Mia’s statements were coded as such. The quotes below exemplify Kim's personal notes when observing lessons with various lenses:

> [Teacher A] chose to allow her students to reach the derivative rule by themselves. This shows how important it is for her to have students try to think and mainly inquire in order to develop their thinking. I connect very much to this idea, and hope to succeed in applying it as a teacher. (Reflection 3.11.21, under the lens of Teacher beliefs)

> [Teacher B] sees importance in repeating everything learned so far about the topic, so that everyone is on the same page and can progress in the current lesson. I’m not sure if this is a belief I agree with. Such recapping can take half of the lesson, and the gain is not necessarily significant. (Reflection 3.11.21, under the lens of Teacher beliefs)

> When checking homework, Teacher B reproached students that did not bring their worksheets. […] They said they forgot, and she said that’s no excuse […]. They responded that she never gives homework, to which she answered that it should be all the more reason to do it in the rare occasion that she does. I did not connect to this response, since it’s the first time […] it is legitimate that they forgot. I was surprised that she reacted so firmly specifically to this issue. (Reflection 3.11.21, under the lens of Interactions)

> [Teacher A] presented the problem on the board and told students to try and solve it on their own, then they will solve it together. After several minutes, she saw that the students are trying and that they have a direction for the solution, and decided to give them time to continue on their own. Eventually, she did not solve the problem in the plenary at all. […] By looking at her students, she probably understood that […] this would be more effective. I hope that as a teacher I would be able to know my students that well, so I would know which is the most efficient way for learning in a specific moment, not only through pre-planning. (Reflection 10.11.21, under the lens of Decision-making)

As revealed in these quotes, in her personal reflections Kim seriously considers whether, and especially why, an observed idea is significant for her to adopt (see the first and the last quotes), reject (third quote), or contemplate on (second quote).

*Development of the two PSTs reflections along the practicum year*. Initially, the requirement to analyze lessons through the lenses posed a significant challenge for Mia, who produced general rather than SLF-based reflections. However, after reading her peers’ reflections on the same lessons, Mia gained clarity on the task and gradually progressed in writing reflections based on the lenses, although she still wrote short blogs and did not include any personal aspects. The quote below is taken from Mia's final summative reflection, in response to the task “look at your submitted reflections

[…] and describe the process you went through, as expressed in these reflections and from your personal perspective”:

> At the beginning of the year, I had no drive […] not for the practicum and not for the reflections. I look at my reflections from the very beginning, and I see this lack of will in my writing, not necessarily the words I wrote but the “atmosphere”, the way I wrote, everything is very concise, I didn't elaborate too much, the lack of energy is very much apparent. Towards the end of the first semester […] I see a change in my reflections there, and I'm really pleased since I also managed to be more focused on the lenses, which is very difficult […]. It's much easier to write when you have a free hand […], but the lenses make it hard, that's their substance. Indeed, I succeeded in eliciting things that I can learn from, and not just frustrations and things that bother me, and I really see the change and it's highly important. In general, I see, and I’m happy about, the process I went through due to the reflections, even though they were annoying. Sometimes when things are written, you simply remember that way.

It appears that Mia genuinely acknowledges her personal growth regarding the ability to extract meaningful insights from lesson observations. It is noteworthy that this recognition occurs despite the fact that at this sum-up point, the requirement to write SLF-based reflections is still seen by Mia as hard and annoying.

Similarly, in her second interview Kim also reveals a view of the lenses as difficult to use, yet points to the concreteness they enable:

> I think that focusing on the lenses was very difficult, but it was also very precise. […] [in any case] I would have said 'wow, it's very nice that I watched the lesson’, but the mere focus on two lenses helped me know what I was looking for, and come out with concrete things. It gave me a really different view of the lesson, for example, when the teacher assigned a task, I immediately thought about why she gave it, what’s the rationale. There was a lesson in the first semester [when we could] watch without analyzing by the lenses, and I remember sitting in that lesson and saying to myself: it’s a really nice lesson, but I don’t have the accuracy. That’s where I felt I need the lenses. It's irritating and frustrating to analyze by the lenses, but it's meaningful. It really provides something. […] I see a lot of meaning in reflections. Something in this whole process was very meaningful.

Kim highlights the lenses' role in providing a structured framework to identify and articulate specific observations. Both Mia and Kim expressed frustration and even irritation about analyzing lessons through SLF. However, whereas Kim acknowledged from the start that such analysis lead to meaningful insights, for Mia this process was longer, with a gradual progress along the practicum year.

## DISCUSSION

In his ‘reflection about reflection’, Russell (2005) maintains that “[in] practicum experiences with those I teach […] Fostering reflective practice requires far more than telling people to reflect and then simply hoping for the best. […] reflective practice can and should be taught - *explicitly*, *directly*, *thoughtfully* and *patiently*” (Russell, 2005, p. 203, emphasis added). Based on this conviction, and encouraged by our successful experiences with developing practicing teachers’ reflective skills using SLF, we embarked on the development of such skills with prospective teachers, before and during their practicum. We did so *explicitly* (in a structured way, namely, by using a specific tool), *directly* (by means of precise prompts), *thoughtfully* (flexibly, that is, allowing teachers to flounder when disconcerted) and patiently (through a long-term period). We were aware that PSTs lack the resources that practicing teachers have at their disposal (e.g., experience and knowledge) and yet we expected that, even if differently, PSTs can develop their own ways to reflect on the mathematics teaching practice. In this study, we asked what may be the foci and characteristics of PSTs' reflective process. The partial set of data we presented in this report allows us to suggest several heartening, even if tentative, answers:

- The development of reflective practices is a slow, at times frustrating, process, which may not be immediately recognized as productive even by outspoken and verbal persons. However, frequent use of the SLF allowed for the development of valued insights which otherwise were unlikely to have been at the forefront of the lesson observations and analysis.
- The use of the SLF further highlighted the differences between the concerns of prospective and practicing teachers (already pointed, for example, by Chamoso et. al, 2012; Lin & Rowland, 2016), and made them explicit. For example, PSTs are preoccupied with general and mathematical pedagogical issues ('how we teach') and with classroom management ('how we organize our teaching'), rather than with the mathematical ideas underlying a lesson topic ('what we teach'), or with mathematically-related teaching dilemmas that are so frequently discussed in many PD sessions with practicing teachers (Karsenty & Arcavi, 2017).
- The focus on revisiting and re-conceptualizing one's own practice, as we witnessed in SLF-based reflections of practicing teachers (e.g., Nurick at al., 2022), is evidently limited in the reflections of PSTs, for whom everything may be novel. We suggest that PSTs, who still lack formed opinions (as well as a practice to reshape), may benefit from how SLF could become a guide for their future profession. Moreover, as the case of Mia demonstrates, collective reflection enhances the process of becoming a reflective practitioner, by learning from others, sharing concerns and exercising exchange of opinions, beliefs and expectations.
- There are individual differences in how reflection practices can be appropriated by different PSTs. Nevertheless, many may benefit, even if at different levels.
- The cumulative effect of watching live lessons of familiar teachers, over an extended period of time, may explain the ultimate recognition (at least by some PSTs) that reflection stirs meaningful insights. This effect can possibly lead to internalizing reflection as a practice to be invoked in order to become an independent critical thinker regarding one's own practice.

These results highlight the importance of introducing reflection as an ongoing, structured and carefully crafted activity in teacher education courses. Beyond the potential benefits for the professionalism of future teachers, it may also be an invaluable tool for teacher educators to know better each of the PSTs under their responsibility, and thus to shape their instruction accordingly.

**Acknowledgments**

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

**References**

Bjuland, R., Cestari, M. L., & Borgersen, H. E. (2012). Professional mathematics teacher identity: Analysis of reflective narratives from discourses and activities. *Journal of Mathematics Teacher Education*, *15*(5), 405-424.

Chamoso, J. M., Cáceres, M. J., & Azcárate, P. (2012). Reflection on the teaching-learning process in the initial training of teachers. Characterization of the issues on which prospective mathematics teachers reflect. *Teaching and Teacher Education*, *28*(2), 154-164.

Chikiwa, S., & Graven, M. (2023). Exploring the development of South African prospective teachers' reflective practice. *Pythagoras*, 44(1), 678.

Finlay, L. (2008). Reflecting on 'Reflective practice'. *Practice-based Professional Learning*, paper 52, 1–27. The Open University. https://oro.open.ac.uk/68945/1/Finlay-(2008)-Reflecting-on-reflective-practice-PBPL-paper-52.pdf

Karsenty, R., & Arcavi, A. (2017). Mathematics, lenses and videotapes: A framework and a language for developing reflective practices of teaching. *Journal of Mathematics Teacher Education*, *20*, 433-455.

Lin, F. L., & Rowland, T. (2016). Prospective and practicing mathematics teachers' knowledge and professional development. In A. Gutiérrez, G. C. Leder, & P. Boero (Eds.), *The second handbook of research on the psychology of mathematics education* (pp. 481-520). Brill.

Nurick, Y., Arcavi, A., & Karsenty, R. (2022). Revisiting the concept of reflection: An empirical investigation of mathematics teachers' articulations on the teaching practice. In J. Hodgen, E. Geraniou, G. Bolondi & F. Ferretti (Eds.), *Proceedings of the 12th Congress of the European Society for Research in Mathematics Education* (pp. 3321-3228).

Rhoads, K., & Weber, K. (2016). Exemplary high school mathematics teachers' reflections on teaching: A situated cognition perspective on content knowledge. *International Journal of Educational Research, 78*, 1-12.

Russell, T. (2005). Can reflective practice be taught? *Reflective Practice*, *6*(2), 199-204. https://doi.org/10.1080/14623940500105833

Schwarts, G., & Karsenty, R. (2020). "Can this happen only in Japan?": Mathematics teachers reflect on a videotaped lesson in a cross-cultural context. *Journal of Mathematics Teacher Education*, *23*, 527–554.

# GENDER DIFFERENCES IN RELATION TO PERCEIVED DIFFICULTY OF A MATHEMATICAL TASK

Bianca Nicchiotti and Camilla Spagnolo

Free University of Bozen-Bolzano

*The theme of perceived difficulty in mathematics is current, but it has only been considered in the last few years. This study aims to highlight gender differences in relation to perceived difficulty of a mathematical task, and factors influencing it. Italy is one of the countries with the largest gender gap in mathematics, hence there is the interest in analysing perceived difficulty considering it as a key to better understand. We started considering as a benchmark the nationwide quantitative analysis of gender gaps in two INVALSI tasks, characterized by different gender gap levels. Then, we link students' perceived difficulty to these two tasks, analysing qualitatively the differences between boys' and girls' perception. Preliminary findings point out that girls' perception is mainly related to personal consideration.*

## INTRODUCTION AND THEORETICAL FRAMEWORK

The topic of difficulty in mathematics has been object of research for years and the difficulty of a mathematical task seems to be influenced by a combination of factors, such as text comprehension (Spagnolo et al., 2021), and affective factors (Zan, et al., 2006). Within this frame, the students' perspective is crucial, and in the last few years research in mathematics education started considering the issue of perceived difficulty (PD) (Spagnolo & Saccoletto, 2023). The PD is different from the difficulty, as the latter is usually evaluated in retrospective, considering the ratio between the students who solved the item correctly and the total number of students who faced the item (Mehrens & Lehmann, 1991).

The issue of perception of difficulty has been analysed in the field of cognitive psychology, developing different characterisations and definitions for the concept (e.g. Eccles & Wigfield, 2020). The one we consider the closest to our conception of PD is the "feeling of difficulty" (FOD), defined as a "metacognitive experience that monitors cognitive processing as it takes place" (Efklides & Touroutoglou, 2010, p. 272). However, FOD and PD are strongly related but conceptually different, as the former is based on the experience whereas the latter can be described as a metacognitive judgement built considering a conscious memory of knowledge about oneself and the task. At the moment, in mathematics education research, there is not a clear definition of PD, but some of the factors that appear to influence it have been described and analysed qualitatively (early phases of the studies were presented by the authors during international conferences ICME14, CERME13, CIEAEM74 and MAVI29). They have been distinguished into five different but not mutually exclusive categories, developed basing on students' answers to specific questions about PD. The first category is *Resolution strategy,* containing all the references to the procedure, strategy or



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 257-264). PME.

knowledge needed to solve the task, calculations and reasonings included. The second one is *Capability and experience,* and it contains elements related to students' opinion of their competence as well as their familiarity with the task and similar previous experiences. The references to the time spent to solve the task, possible obstacles encountered and doubts belong to this category too. The third one, *Emotions*, regards the emotional aspects (positive or negative). The fourth one is *Task formulation*, including all the comments about the task in general (structure, text, type, etc.) while the last one, *Personal consideration*, concerns each student's personal opinion about their own success in mathematics or as a student.

In this paper we want to analyse students' PD of a task from a gender perspective. The issue of gender gap in mathematics is well-known and current; in fact national and international surveys state that, in many countries, boys achieve better results than girls in mathematics at all school levels (Giberti, 2019); moreover, Italy is one of the countries with the largest gender gap in favour of boys (OECD, 2019). Some Italian studies confirmed that the cause of gender gap cannot be found in any biological or cognitive difference between boys and girls; instead, a variety of factors can play a role in it, such as metacognitive influences, affective factors and general biases (Giberti et al., 2016). There is the urge to put in place didactical interventions to equally involve boys and girls in mathematics, aiming to a more equitable discipline (Ferrara et al., 2021). To do so, we consider crucial to deepen the description of the phenomenon also considering students' perspective. When solving tasks, boys and girls seem to have different approaches; not only usually more boys than girls solve the tasks correctly but also, among the wrong answers, the two groups appear to prefer different distractors (Giberti et al., 2016). Hence, the aim of the paper is to discuss students' PD of tasks, highlighting whether there are any differences between boys' and girls' perception or not, even in relation to the categories of factors influencing it.

## METHODOLOGY

We built a questionnaire composed by two INVALSI tasks, administered in previous years to grade 10 students, each one followed by a set of questions. The selected INVALSI tasks, represented in Figure 1, are both argumentative and dealing with algebra, but they are of a different type, in fact Task 1 is a multiple-choice question while Task 2 is an open-ended one.

**Task 1**

Antonio states that «4n-1 is always a multiple of 3».
Is Antonio right?
In the following table select the only argumentation that justifies the right answer

| Antonio is right... | Antonio is not right... |
|---|---|
| A. ☐ because 4n-1=3n | C. ☐ because 4n-1 is always odd |
| B. ☐ because if n=4 then 4n-1=15 | D. ☐ because if n=3 then 4n-1=11 |

**Task 2**

Marco states that, for every natural number n greater than 0, $n^2+n+1$ is a prime number.
Is Marco right?

Choose one of the two answers and complete the sentence.

☐ Marco is right because ..........................................

☐ Marco is not right because ..........................................

Figure 1: Task 1 and Task 2 (original texts from www.gestinv.it., translation provided by the authors).

In particular, Task 1 asks the students to select the argumentation supporting the right answer, whereas Task 2 requires the students to construct and provide themselves an argumentation for a given statement.

The tasks selected were quantitatively analysed and used as benchmarks. From the national data, 40.3% of Italian students answer correctly to Task 1 and 17.8% respectively to Task 2, meaning that they have a different difficulty in the traditional sense of the term. Furthermore, Task 1 and 2 were afflicted by gender gap in different proportion. For each one, we calculated the value of the gender gap index $GGI_k$ (Spagnolo & Nicchiotti, 2023), defined as follows

$$GGI_k = \begin{cases} \frac{M_k - F_k}{F_k}, \ M_k > F_k \\ \frac{M_k - F_k}{M_k}, \ M_k < F_k \end{cases} ,$$

were $M_k$ is the ratio between the number of the correct answers to the item given by boys and the total number of answers to the item given by boys, while $F_k$ is the equivalent for girls. The value of the index is positive when boys overperform girls, negative in the opposite case and equal to 0 when there is no gender gap. Referring to the threshold values provided, Task 1 resulted to be balanced ($GGI_k = 1.4\%$) while Task 2 has a moderate gender gap in favour of boys ($GGI_k = 16.5\%$).

In the questionnaire, students were asked to solve the tasks; each task was followed by the request for students to rate it according to their PD, on a scale from 1 to 10 (being 1 “very easy” and 10 “very difficult”) and from the request to motivate their rating. The last part of the questionnaire asked students which task was the more difficult in their opinion and why; these last questions gave us more elements to analyse their PD in general. After these preliminary analyses, we carried out a qualitative study involving 7 classes from two Italian high schools (5 grade 9 and 2 grade 10 classes) for a total of 148 students, of which 61 boys and 87 girls. The two schools were a technical scientific high school (Istituto tecnico in Italian) and an educational humanistic high school (Liceo delle Scienze Umane in Italian); the classes from the former had a predominance of boys, while from the latter had predominantly girls. The questionnaire was administered through Google Forms and students answered it during class hours using school computers. The answers were then collected and analysed qualitatively, carrying out a text analysis on students’ answers referring to the categories of PD previously discussed. In addition to that, mean values of the students’ ratings of difficulty were calculated as well as the $GGI_k$ referring to the results of the students of the sample. All the above was done considering the gender perspective and highlighting possible similarities and differences.

## RESULTS AND DISCUSSION

Firstly, quantitative analysis allowed us to select Task 1 and 2 (Figure 1), since they are meaningful from a gender gap perspective. Then, we analysed qualitatively the answers of 148 students to the two tasks (determining whether they had solved them correctly or not) and we calculated the value of the gender gap index for each task referring to the results of our sample. In contrast to the values obtained for the national data, in this case we observed for both tasks the presence of a severe gender gap in favour of males. In fact, the value of the index for Task 1 (T1) is the same as for Task 2 (T2), namely $GGI_{T1} = GGI_{T2} = 41\%$; $GGI_{T1} = GGI_{T2}$ by coincidence, but the presence of a gender gap in favour of males clearly emerges for the two tasks, considering our sample. We consider important to treat this as a characteristic of our sample when deepening the analysis of PD and the categories of factors; to be statistically significant the study should be implemented quantitatively.

### Analysis of PD for Task 1

Students regarded Task 1 as moderately difficult, with the girls seeing it on average more difficult than the boys. The average difficulty assigned by the girls is, in fact, equal to 5.5 while for the boys it is 4. Therefore, this result seems consistent with the presence of a gender gap highlighted by the value of $GGI_{T1}$: not only girls performed more poorly on this task, but they also found it more difficult than the boys did.

The textual analysis of the reasons of the rating allowed us to outline a more precise picture of the issue. We identified and classified 171 references, meaning that some answers presented elements recalling more than one category. 39 of them were not considered in the analysis as they did not give any information, being answers like “It was difficult” or left blank. The remaining ones (59 from boys and 73 from girls) were classified into the five categories, as summarized in Table 1.

| Category | Number of references (boys) | Number of references (girls) |
|---|---|---|
| Resolution strategy | 28 | 35 |
| Capabilities and experience | 24 | 15 |
| Emotions | 2 | 1 |
| Text formulation | 5 | 10 |
| Personal consideration | 0 | 12 |

Table 1: Distribution of the references among the categories for Task 1.

For both boys and girls, the broadest category resulted *Resolution strategy* with many references to the reasoning for boys and to examples for girls. However, for boys, this category seems to have almost the same importance as the category of *Capabilities and experience*, which for girls too is the second broadest but not with a comparable importance. Hence, most students focused on the resolution of the task as the crucial element to determine the reasons of their PD, considering primarily the objective aspects and leaving aside the more subjective side. This might be an explanation for the almost total absence of any reference to *Emotions*: students might not consider this factor when rating the difficulty of a task because they are used to leave out the subjectivity when dealing with mathematics and they might think about the emotional side only if directly asked about it. The category of *Text formulation*, for this task, appears to play a role in the PD for both boys and girls but it is not the most significant factor. Almost all the references belonging to this category mentioned the text while only a few ones referred to the question format.

The most evident difference between boys' and girls' answers is represented by the category of *Personal consideration*. As already described, it contains the answers mentioning students' consideration of themselves under different points of view: them as students, their relationship with mathematics, their reflection about their progresses in their course of study and others. For Task 1, elements recalling these aspects were mentioned only by girls. None of the boys considered this kind of factors important when determining their PD; girls instead considered them even more important than the text formulation. They reflected about their self-perception ("I do not have logic"; "I do not understand these things and I cannot do them even though I practice a lot"), their perception as mathematics students ("I am not good at maths"; "Probably I am not very good at solving these tasks") and their preparation ("I should practice more"; "I should study better").

**Analysis of PD for Task 2**

Like Task 1, Task 2 was considered of a medium difficulty but in general slightly more difficult than the former. The average rating given by boys is equal to 4.1 while the one from the girls is 5.7. In this case, the difference between the average ratings is as evident as before and it confirms the findings described for Task 1: girls tend, on average, to perceive tasks as more difficult than the boys do. Despite Task 2 resulting very difficult, as national data show, the task was perceived as only moderately difficult. However, we observed that in this case, students' perception was not aligned with their actual results, in fact the percentages of right answers were even lower than the national ones (11% of the students answered correctly, 9% of the girls and 13% of the boys). Students seem to lack awareness of their mistakes and they are not always coherent rating the task and explaining the reasons, almost as if they are worried about giving a too high rating to the PD of the task.

Analysing students' answers and explanations we identified 177 references, 39 of which were not meaningful. The remaining 138 (59 from boys' answers and 79 from girls' ones) were classified into the five discussed categories, as reported into Table 2.

| Category | Number of references (boys) | Number of references (girls) |
|---|---|---|
| Resolution strategy | 27 | 35 |

| | | |
|---|---|---|
| Capabilities and experience | 28 | 22 |
| Emotions | 0 | 0 |
| Text formulation | 4 | 13 |
| Personal consideration | 0 | 9 |

Table 2: Distribution of the references among the categories for Task 2.

The results for Task 2 are similar to the ones obtained for Task 1, making the findings stronger. For girls, once again, the most represented category is *Resolution strategy* with references to examples needed to solve the task, reasonings that can be put in place and calculations. In particular, this last element was one of the biggest concerns for girls, as they considered difficult dealing with the powers. For boys too this category is one of the most important ones, as almost all the references are divided evenly between *Resolution strategy* and *Capabilities and experience*. In the former, boys did not refer to calculations as much as girls did; they instead focused on the examples and the reasoning. The reasoning in particular has both a positive and a negative connotation for boys, because some of them stated that “more reasoning was needed to solve the task” while others wrote that “the task required not a lot of reasoning”, being this last one the prevalent thought. Regarding the category of *Capabilities and experience*, it is well represented both for boys and girls, being the most important one for the former. Boys mentioned very often in their answers their confidence in the answer provided and the absence of obstacles encountered, as a reason for considering the item easy or very easy. From the answers belonging to this category, we can observe the absence of awareness of mistakes mentioned before, especially from boys. The girls’ answers from this category contain many references to obstacles but considered only in the negative sense. In other words, girls considered the item more difficult because they dealt with many different obstacles solving it, and they seem more aware of this aspect. The other factor considered by them is the previous experience with similar tasks, because girls are quite honest about the fact that they consider a task more difficult if they never solved something similar. Once again, we could not find any reference belonging to the category *Emotions*. For Task 2 as well as for Task 1, the category *Text formulation* results not to be the most important one. In this case, boys’ and girls’ answers are similar, considering confusing above all the presence of “many letters in the text” of which they do not know the values. Many students affirmed that the task would have been easier if in the text there had been some “example of numbers to substitute to n”.

Finally, the answers belonging to the category of *Personal consideration* are again the element that mainly differentiates boys’ and girls’ answers. In fact, among boys’ answers, also for Task 2 there were not references to this category, while girls mentioned elements related to it. In this case, the category is slightly smaller than in Task 1, but the previously described characteristics are preserved. Girls seem more conscious about their mistakes and difficulties, sometimes even making general and hard remarks about themselves, which do not seem to descend directly from the task itself (“I cannot do it, I have basic gaps”, “I am ignorant”). An interesting aspect to evaluate is the fact that the girls whose answers can be classified into this category for Task 2, referred to this category also in their answers regarding Task 1. Moreover, in most cases, they considered the two tasks equally difficult and they attributed very high ratings to them. This evidence might suggest that personal consideration and PD are intertwined for girls and a negative personal consideration could be both a cause and an effect of a high PD of a task.

## CONCLUDING REMARKS

The study presented in this paper gives a first insight into the PD and the factors characterizing it, comparing boys’ and girls’ perspectives and analysing them on the basis of the issue of gender gap in mathematics. The analysis of students’ ratings and answers allowed us to state that there seem to be some differences between boys’ and girls’ PD. Boys tend to evaluate mathematical tasks as easier than girls do, even regardless of their actual performance solving them.

The categorization of the references into the five sets allowed us firstly to confirm their usefulness to describe the factors influencing students’ PD. Moreover, we observed that the important categories to determine PD seem to be almost the same for boys and girls. The main difference is represented by the connotation they give to the elements: the same aspect, in fact, is ambivalent for boys, being both positive and negative, and usually negative for girls. Namely, when motivating their PD, girls tend to be very severe highlighting all the aspects that make a task more (or less) difficult, but they almost never consider the factors making it easy. This implies that the same element such as the text of the task, might result in a facilitating element for boys and a distractor for girls. However, the biggest difference between boys and girls regarding the categories is the complete absence for boys only of any reference to the *Personal consideration*. It is one of the categories involving subjectivity, which instead appears to be discarded by boys when elaborating about their PD. On the other hand, girls always make reference to it and although it is not the most important category for them either, it is always considered. Element pertaining to it are expressed especially by the girls that had difficulties or were not able to solve the tasks. They appear to have a very low personal consideration and make hard judgement about themselves not only as students but as people in general. This led us to hypothesize the existence of a cause-effect link between PD and personal consideration for girls that works both ways. More studies are needed to confirm this, but it is possible that girls’ low personal consideration makes them perceive mathematical tasks as more difficult, and it seems also reasonable that perceiving mathematical tasks as very difficult leads them to build a bad self-opinion resulting in low personal consideration.

The validity of these findings could be strengthened considering expanding the study working with a bigger sample and more tasks, of different types and regarding other topics than algebra. This is important with the prospect of understanding the reasons behind the existence of a gender gap in mathematics and, above all, to put in place concrete actions to fill the gap and make mathematics fairer.

## References


Eccles, J. S., & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology, 61*. Art 101859. https://doi.org/10.1016/j.cedpsych.2020.101859

Efklides, A., & Touroutoglou, A. (2010). Cognitive interruption as an object of metacognitive monitoring: Feeling of difficulty and surprise. In: A. E. &. P. Misailidi (Ed.), *Trends and prospects in metacognition research* (pp. 171–208). New York, NY: Springer. https://doi.org/10.1007/978-1-4419-6546-2_6

Ferrara, F., Ferrari, G., Robutti, O., Contini, D. & Di Tommaso, M.L. (2021). When gender matters: A study of gender differences in mathematics. In Inprasitha, M., Changsri, N., & Boonsena, N. (Eds), *Proc. 44th Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 2, pp. 255-263). PME.

Giberti, C., Zivelonghi, A. & Bolondi, G. (2016). Gender differences and didactic contract: Analysis of two INVALSI tasks on power properties. In Csíkos, C., Rausch, A., & Szitányi, J. (Eds), *Proc. 40th Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 2, pp. 275-282). PME.

Giberti, C. (2019). Differenze di genere in matematica: dagli studi internazionali alla situazione italiana. *Didattica della matematica. Dalla ricerca alle pratiche d'aula*, 44-69.

Mehrens, W. A., & Lehmann, I. J. (1991). *Measurement and evaluation in education and psychology*. Fort Worth: Holt, Rinehart & Winston.

OECD. (2019). *PISA 2018 Results (Volume II): Where All Students Can Succeed*. Paris: OECD Publishing. https://doi.org/10.1787/b5fd1b8f-en

Spagnolo, C., Capone, R., & Gambini, A. (2021). Where do students focus their attention on solving mathematical tasks? An eye tracker explorative study. In Inprasitha, M., Changsri, N., & Boonsena, N. (Eds), *Proc. 44th Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 4, pp. 84-92). PME.

Spagnolo, C. & Nicchiotti, B. (2023). Interpreting gender gap issues in standardized tests: definition and application of a theoretical tool. *Frontiers in Education, 8.* https://doi.org/10.3389/feduc.2023.1303041

Spagnolo, C., & Saccoletto, S. (2023). How students view the difficulty of mathematical tasks: factors that influence their perceptions. In P. Drijvers, C. Csapodi, H. Palmér, K. Gosztonyi, & E. Kónya (Eds.), *Proceedings of the Thirteenth Congress of the European Society for Research in Mathematics Education (CERME13)* (pp. 1498–1506). Alfréd Rényi Institute of Mathematics and ERME.

Zan, R., Brown, L., Evans, J., & Hannula, M. (2006). Affect in mathematics education: an introduction. *Educational studies in mathematics, 63*(2), 113–121. https://doi.org/10.1007/s10649-006-9028-2

# PROSPECTIVE TEACHERS' UNDERSTANDING OF THE INDIRECT PROOF OF THE CONVERSE OF THE INSCRIBED ANGLE THEOREM

Masakazu Okazaki and Keiko Watanabe

Okayama University and Shiga University, Japan

*This study clarifies the geometric thinking of prospective teachers on the converse of the inscribed angle theorem toward van Hiele's fourth level concerned with the understanding of an indirect proof structure. As a result of the analysis, the understanding process was divided into four stages from two perspectives: acceptance of the theorem and construction of the indirect proof. The difference between the first and second stages relates to whether the background theory is the direct proof scheme or logic that permits the intuitive acceptance of the theorem. The third stage relates to indirect argumentations that are facilitated by assuming impossible objects. The final formal proof stage relates to the logical structure under conditions for which the secondary statement involving impossible objects can be proved as contradictory.*

## INTRODUCTION

Proofs not only express the validity of mathematical propositions but also reflect our process of understanding the propositions (Hanna, 2008; Boero, 2012). However, the meaning of a geometric indirect proof is not directly connected to the geometric phenomenon of the proposition, as the proof assumes a world in which the statement does not hold or it uses the contrapositive of the statement. From this point of view, great difficulties arise in generating and/or understanding indirect proofs (Antonini, 2019). However, prospective teachers must understand indirect proofs to teach mathematics. In one example, Japanese ninth-grade textbooks deal with the converse of the inscribed angle theorem (CIAT) and provide an explanation of the CIAT, albeit not a formal proof related to indirect proofs.

The aim of this study is to determine what level of geometric thinking prospective secondary-school mathematics teachers display regarding their understanding of indirect proofs. We build on van Hiele's levels of thinking (incorporating levels 1 to 5 in his model) to determine student teachers' understanding of a geometric proof, from the third level to the fourth level, which concerns the understanding of the structure of a proof. Van Hiele (1985, p. 48) describes the third level as follows: "Properties are ordered. At this level the intrinsic meaning of deduction is not understood by the students". He describes the fourth level as follows: "Thinking is concerned with the meaning of deduction, with the converse of a theorem, with axioms, with necessary and sufficient conditions". It is necessary to use the thinking of the fourth level to understand the indirect proof (van Hiele, 1986). Numerous studies on proofs and proving have focused on how students achieve the third level of thinking, but only a few studies have investigated students' achievements toward the fourth level (Battista,



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 265-272). PME.

2007). The initial transitional stage from the third to fourth level has been examined as a process of conceptualizing a proof in terms of a symbolic linguistic description. The present study clarifies how prospective teachers display their geometric thinking toward van Hiele 's fourth level by analyzing their construction and explanation of the proof of the CIAT.

## THEORETICAL BACKGROUND

Antonini and Mariotti (2008) argued that a mathematical theorem comprises the triplet (S, P, T) of a statement (S), proof (P), and theory (T), and that the third component, the theory, represents both a mathematical theory such as Euclidean geometry and a logic that represents rules of inference. Furthermore, they argued that in an indirect proof, the principal statement S must be transformed once into a secondary statement S* and that the transformations between S and S* depend on a classic logic. They then argued that the indirect proof is a meta-theorem (MS, MP, MT) (which comprises a meta-statement MS = $S^* \rightarrow S$ and a meta-proof MP based on a specific meta-theory MT (logic)) with sub-theorems (S*, C, T)) (Antonini and Mariotti, 2008, p. 405). In indirect proofs, the elements applied by a meta-theory are the secondary statements (S*), and furthermore, the direct proof of S* deals with "impossible" objects; e.g., when proving if A then B, one assumes an object that is A and not B in the proof by contradiction. The difficulty of the indirect proof depends on whether this impossible object can be considered and whether the theory of logic can be applied to it.

In addition, Antonini (2019) explained that intuitive acceptance operates in generating proofs by contradiction according to the theory of figural concepts (Fischbein, 1982), and he showed that indirect argumentation can be generated in a compromising manner as the oscillation between figural and conceptual elements. Here, argumentation is regarded as "whatever rhetoric means are employed in order to convince somebody of the truth or the falsehood of a particular statement" (Antonini, 2019) and the process from argumentation to proof is considered from the perspective of cognitive unity. Moreover, Garuti et al. (1998) applied the theory of cognitive unity to the process of theorem exploration and proof generation. We consider that both conceptions of cognitive unity are important in analyzing how students construct indirect proofs. Meanwhile, it has been pointed out that there is a need to focus more on the content area with regard to the process of constructing indirect proofs (Dawkins et al, 2016; Hakamata et al., 2023), and we thus conduct a detailed analysis of students' mathematical understanding of the CIAT.

## METHODS

### Proof of the CIAT

The method of proof by conversion is described as follows (Kodaira, 2015). In general, a proposition is expressed as 'if $\Gamma$ then $\Delta$', and three theorems, namely 'if $\Gamma_1$ then $\Delta_1$ ', 'if $\Gamma_2$ then $\Delta_2$', and ' if $\Gamma_3$ then $\Delta_3$', are assumed to be true. If the assumptions $\Gamma_1$, $\Gamma_2$, and $\Gamma_3$ are satisfied in all cases and no two of the three conclusions $\Delta_1$, $\Delta_2$, and $\Delta_3$ are incompatible, then the converses of the three theorems, namely 'if $\Delta_1$ then $\Gamma_1$', 'if $\Delta_2$

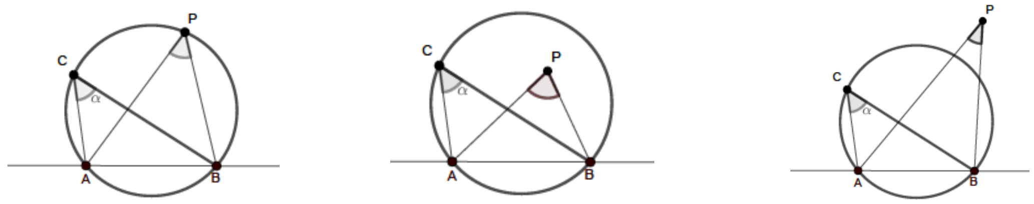
then $\Gamma_2$', and 'if $\Delta_3$ then $\Gamma_3$', are all true. The principle of proof by conversion is based on proof by contradiction: if the statement ‘if $\Delta_1$ then not $\Gamma_1$’ is assumed to be true, then it must be that $\Gamma_2$ or $\Gamma_3$, because all cases are satisfied by $\Gamma_1$, $\Gamma_2$, and $\Gamma_3$. At this point, a contradiction arises because it has already been proved that 'if $\Gamma_2$ then $\Delta_2$' and 'if $\Gamma_3$ then $\Delta_3$', and that $\Delta_2$ and $\Delta_3$ are also incompatible. Hence, if $\Delta_1$ then $\Gamma_1$.

The CIAT is stated as follows. If the two points C and P are on the same side with respect to the line AB, then ∠APB = ∠ACB. In Japan, the CIAT is taught in the ninth grade using the textbook description shown in Figure 1.

[1] If point P is on the circumference of the circle, then ∠APB = ∠$\alpha$

[2] If point P is inside the circle, then ∠APB > ∠

[3] If point P is outside the circle, then ∠APB < ∠$\alpha$




We thus see that ∠APB = ∠α only if point P is on the circumference of the circle.

Figure 1 Textbook description (explanation A)

It is confirmed first that (1) "if point P is on the circumference of the circle, then ∠APB = ∠α", (2) "if point P is inside the circle, then ∠APB > ∠α", and (3) "if point P is outside the circle, then ∠APB < ∠α" and then that "∠APB = ∠α is true if point P is APB = ∠α is only true if point P is on the circumference of the circle" (Figure 1, labelled explanation A). The CIAT is then formulated as “When two points C and P are on the same side of the line AB, if ∠APB = ∠ACB, then the four points ABCP lie on one circumference”.

Figure 1 shows part of the proof that guarantees this theorem. The assumption and conclusion are vice versa in case [1] (the inscribed angle theorem) and the formulation of the CIAT. The analysis in the present paper focuses on how students pay attention to this point for acceptance of the theorem and how they understand this explanation in constructing a proof of the conversion method.

**Interviews and analysis**

An interview survey was conducted to obtain the prospective teachers’ understanding of the CIAT. The subjects were 11 fourth-year students majoring in secondary-school mathematics education and belonging to two national universities in Japan. Each interview was conducted for a pair of students (except for one interview conducted for a sole student) and took approximately 1 hour. The students were allowed to talk freely with each other.

First, the contents of the inscribed angle theorem presented in the textbook for the ninth grade at junior high school were reviewed with the interviewer. Next, the interviewer confirmed with the students that the CIAT is described after explanation A in the textbook. The interviewer first checked to see whether the students could find a reverse relation between the assumptions and the conclusion and then asked the students why the CIAT could be formulated by presenting explanation A. This question was asked to reveal what kind of argumentation or proof the students would construct.

The interview data were transcribed. The students' ideas and explanations were then encoded, qualitatively interpreted, and conceptualized in terms of how students made sense of the theorems and constructed proofs. In particular, we focus on the types and aspects of argumentation, situations in which intuitive acceptance comes into play, the interaction between a diagram and reasoning, and the interaction between the exploration of theorems and the generation of proofs.

## RESULTS

First, when the interviewer asked whether there was any discrepancy when comparing case [1] in explanation A with the statement of the CIAT, all students noted that the assumption and conclusion were reversed. However, when asked why they could formulate the CIAT on the basis of explanation of the three cases in Figure 1 (i.e., explanation A), distinct answers were given for the acceptance of the theorem and the organization of the proofs. In the following, we examine the results of the interviews on the three cases in explanation A.

### Case of students Ak and Km

In the interview with the pair of students Ak and Km, Ak demonstrated an understanding of the indirect proof from the beginning whereas Km was unclear whether the theorem was even true.

Km: I don't think we have proven yet that point P is on the same circumference if APB = $\angle\alpha$ = ACB.

Ak: No, no, if you just look at case 1, you can't state the correctness, but you can state it if you look at cases 1, 2, and 3.

Km: I could, but, you know, what can I say, it's too much of a jump in logic. Because nothing is written between them (explanation A and the CIAT).

Ak: So, you know, point P is not on the circumference. We therefore have case 2 or 3. You know, the angles are not equal then, as shown by the case separation. So, (junior secondary students) don't know the contrapositive. Maybe you can't reverse the assumption and conclusion. I think what you're doing is proving the contrapositive.

Ak insisted that explanation A can be seen as the contrapositive, and he claimed that the explanation in the textbook is the basis for proving the theorem, whereas Km continued to insist that there is a jump in logic for the reverse of the assumption and conclusion. The conversation between the two proceeded without one understanding the other. Finally, Ak offered a clearer explanation.

Ak: ... If the four points ABCP are not on one circumference, then they are inside or outside. Here, it is case 2 or 3. Because it is smaller or larger than APB. So that means ∠APB does not equal ∠ACB. It probably won't be obvious to most junior high school students. But, I think that's the only way to explain it.

Km, in contrast, argued that she wanted to create a direct proof, saying that "when considering proofs for junior high school students, it is better that the proof be direct, without the use of the contrapositive and so forth". Km then insisted the method of checking whether the centers of two circumscribed circles coincide.

Km: Yes, ABP. One circle is determined. ACB is the circumferential angle to AB, so there is a center point. ... If you can say the same thing with a circle at ABP, and the center point is determined. If the two center points are the same, then point P is determined to be on the same circumference. If the two center points deviate, it cannot be said that the points are on the same circumference. Do you think it good the idea that ABC's circle and ABP's circle are exactly overlapping?

Ak was clearly aware of the statement S and transformed it into the secondary statement S*. In contrast, Km felt an inconsistency in that the assumption and conclusion between case 1 and the CIAT was reversed, and she tried to think of the theorem's assertion in a direct proof using a diagram. The background theory for Ak was logic that could deal with the contrapositive, whereas that for Km was Euclidean geometry enabling a direct proof.

**Case of students Ms and Ky**

Students Ms and Ky were convinced that the theorem held true by explanation A. However, Ms, despite mentioning the indirect proof, was not conscious that she had created the proof.

Interviewer: After checking this explanation, can we say this theorem is true?

Ms: I'm comfortable that the theorem holds.... because it's only if P is circumscribed. If the angles are equal, we know point P is circumscribed. If it's just this case (case [1]), they don't connect, but because there are three cases, I think the assertion holds true.

The interviewer tried to introduce a conflict, but the students were unfazed.

Ms: It can be said that if point P is *not on the circumference, then it's not equal*. If it is not on the circumference, then the angles are not equal… If the angles are equal, then the points are on the circumference.

Ky: If we examine all of the circumferential, internal, and external cases, angle APB is greater than angle α in the internal case and less than angle α in the external case. In the circumference case, they are equal. So, if the angles are equal, then ABCP is on one circumference. ... It's the only way. I can't explain it any further.

Student Ms made a further argument that "As it is the case of only if, conversely, if ∠APB = ∠ACB, then it is on the circumference." However, she didn't give the formal explanation of the indirect proof. When the interviewer gave the formal description of the contrapositive in the last stage of the interview, Ms reflected on her reasoning as follows.

Ms: I wondered whether the points are absolutely on the circumference when the angles are equal. When I wondered if there are cases that it is not on the circumference, I didn't notice that such cases can be proved using cases 2 and 3.

Both Ms and Ky were able to accept the theorem, but they were at different levels in constructing the indirect proof. Ky was conscious of the correctness of the theorem but did not give the argumentation. Ms intuitively grasped the structure of the indirect proof but could not use cases 2 and 3 in explanation A to describe the proof. We thus think that Ms had not yet reached the level of a formal proof.

**Case of students St and Ny**

The most common occurrence was that the students felt an inconsistency in the reversal of the assumption and conclusion but did not fully understand the CIAT and attempted to create a direct proof. Students St and Ny raised doubts about the generality of the explanation in the formulation of the proposition, and moreover, the proof that they attempted was a graphic search in which they drew auxiliary lines on the diagram.

St: There are four points, and there is a center equal distant from the four points. There is a point from three points. I wish to show that another circle is also equal.

Ny: I wish to produce a point O at the central angle. Then, the circle of APB and the circle of ACB become the same.

A common trend for other students who did not do well was that they thought the two circles of APB and ACB were congruent.

## DISCUSSION

We consider that the two perspectives of acceptance of the theorem and the composition of the indirect proof enable an explanation of the prospective students' understanding of an indirect proof as a transitional stage from van Hiele's third level to fourth level. Just 3 of the 11 students who had already reached the third level seemed to achieve the fourth level, or semi-proficiency (Stages 3 and 4 in Table 1).

| | | Acceptance of the theorem | Construction of the indirect proof |
|---|---|---|---|
| Stage 4 | Ak, Nn | Yes | Yes |
| Stage 3 | Ms | Yes | Yes (intuitively) |
| Stage 2 | Mc, Ky | Yes | No |

| | | | |
|---|---|---|---|
| Stage 1 | Km, St, Ny, Mm, Ym, Yr | No | No |

Table 1: Stages of accepting the theorem and constructing the indirect proof

Stage 1 is a stage of thinking in which the principal statement S is used directly in the proof relying on a diagram in a Euclidean way, as in the case of Km. As the properties implied by the principal statement S are explored through the diagram, the principal statement is not regarded as the starting statement for an indirect proof; i.e., the secondary statement S* is not generated. The tendency to fall into diagrammatic inquiry has been noted by Dawkins et al. (2016) and Antonini (2019). Stage 2 is the stage in which, as in the case of Ky, the consistency between the principal statement S and the given explanation A for it, namely what three cases in explanation A guarantee the validity of the CIAT, is intuitively recognized, but no argumentations for it are created. In this stage, the secondary statement S* as the start of the indirect proof is not expressed at the logical level, as assuming the "impossible" object is not recognized as a crucial element of the logical structure for proving the principal statement S. In Stage 3, a secondary statement S* is generated from the principal statement S at the conceptual level, and indirect argumentations are constructed by making sense of the three cases, but it is not recognized that the given cases can be used as part of a formal indirect proof, as in the case of Ms. Here, an indirect argumentation remains intuitive and unconscious as a formal proof. In Stage 4, as in the case of Ak, the principal statement S is conceived as a starting proposition for the indirect proof, from which the secondary statement is constructed and the indirect proof generated. The students understand that the given explanations guarantee that the contradictions emerge from the secondary statement.

To move toward indirect proofs, it is necessary to shift the background theory from direct proofs according to Euclidean opportunities to classical logic, and yet several stages were observed in the process of constructing indirect proofs, especially those recognized by students in Stage 3. Ms stated that "I thought about whether the points are absolutely on the circumference when the angles are equal". We think that her idea of "whether something absolutely exists or not" is important, because it may naturally lead to the idea of constructing the secondary statement S* from the principal statement S and to explore whether the secondary statement is true or not. In this sense, the student could think of the impossible object in terms of the $S \rightarrow S^*$ relation and intuitively construct the indirect argumentations. However, to reach Stage 4, it is important for the student (such as in the case of Ak) to recognize the logical structure in which the three cases exhaust all cases and include cases in which the secondary statement involving impossible objects can be proved as contradictory, which enables an indirect formal proof.

In this study we explored the transition from van Hiele's third level to fourth level on the basis of proof by conversion. We consider proof by conversion a suitable example with which to clarify students' attainment of the fourth level because of the difficulties involved in understanding such a proof for it includes the ideas of the other indirect proofs. However, we need to consider other content areas, such as algebra, to further clarify the logical development.

**Acknowledgment**

This work was supported by JSPS KAKENHI Grant Numbers 20H01745 and 23K02409.

**References**

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

Antonini, S. (2019). Intuitive acceptance of proof by contradiction. *ZDM Mathematics Education*, *51*, 793-806.

Battista, M. (2007). The development of geometric and spatial thinking. in F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 843-908). National Council of Teachers of Mathematics.

Boero, P. (2007). *Theorems in school: From history, epistemology and cognition to classroom practice*. Rotterdam: Sense Publishers.

Dawkins, P. and Karunakaran, S. (2016). Why research on proof-oriented mathematical behavior should attend to the role of particular mathematical content. *The Journal of Mathematical Behavior, 44*, 65-75.

Garuti, R., Boero, P., & Lemut, E. (1998). Cognitive unity of theorems and difficulty of proof. in A. Olivier *et al*. (Eds.), *Proc. of the 22nd Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 2, pp. 345-352). PME.

Hakamata, R., Uegatani, Y., and Hayata, T. (2023). How do open-ended problems promote indirect argumentation? Conceptual replications in a Japanese secondary school. *The Journal of Mathematical Behavior*, *70*, 65-75.

Hanna, G., & de Villiers, M. (2012). *Proof and proving in mathematics education. The 19th ICMI study*. Dordrecht: Springer.

Kodaira, K. (2015). *The joy of geometry*. Iwanami shoten.

van Hiele, P. (1985). The child's thought and geometry. Fuy, D. et al.(Eds.), *An investigation of the van Hiele model of thinking in geometry among adolescents* (pp.243- 252). Brooklyn, NY: Brooklyn College.

Van Hiele, P. (1986). *Structure and insight: a theory of mathematics education*. Orlando: Academic Press.

# THE SIGNIFICANCE OF TEACHING TO RECOGNISE THE MATHEMATICAL TERMS AND NOTATIONS

Lucian Olteanu

Linnaeus University, Department of Mathematics, Sweden

*While the concept of a function has been extensively researched worldwide, there has been limited investigation into how functions are taught in the classroom and the students' opportunities to understand the uses of notation in two upper-secondary classroom settings. This paper aims to address this gap by examining the content presented in a textbook and within the teaching that occurs in two upper secondary classrooms. Data for the analysis were collected from two classes, involving two teachers and 45 students, and included video recordings of lessons and tests. The analytical framework is grounded in variation theory. The findings underscore the crucial role of teaching in providing students with the opportunity to discern the meaning of the notation related to the concept of a function.*

## INTRODUCTION

The concept of function is of great importance in mathematics education. However, numerous researchers have documented various misconceptions and difficulties experienced by students concerning this concept. For instance, research has reported students' misconceptions regarding interpretation or meaning, graphic representation, and function characteristics (e.g., Chen, 2023; Dogan-Dunlap, 2007; Thompson & Carlson, 2017). Furthermore, research has highlighted students' learning difficulties related to the concept of functions, encompassing issues with the definition of functions, interpretation or meaning, as well as notation and expression (e.g., Clement, 2000; Sajka, 2003). Musgrave and Thompson (2014) as well as Thompson and Milner (2019) have redirected their attention from students' comprehension of function notation to that of teachers. According to their findings, many teachers perceive function notation primarily as a label or a name for the defining formula, rather than as a representation of the values of one quantity with another. However, few studies focus on the teaching that occurs in the classroom, especially with a focus on notation in upper secondary school (Minh & Lagrange, 2016; Olteanu & Olteanu, 2012). Further research is still needed since current students continue to have learning difficulties with the concept of function (Trujillo et al., 2023). The present article may contribute to the development of this research by analysing the treatment of functions in two upper-secondary classrooms and the opportunity that students have to understand the uses of notation connected to the concept of function.

In functions that rely on a single variable, the convention is to use $x$ to represent an element from the domain, and $f(x)$ to denote the associated element in the range. $x$ is termed an independent variable since it pertains to domain elements that select elements from the range. The connection between x and $f(x)$ can be expressed



2024. In T. Evans, O. Marmur, J. Hunter, G. Leach, & J. Jhagroo (Eds.). *Proceedings of the 47th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 273-280). PME.

algebraically through a specific rule. Sometimes, the elements of a range are denoted as $y$, which usually leads to confusion between equation and function (e.g., Olteanu & Olteanu, 2012; Trujillo et al., 2023). $x$ is termed a variable because it has the capacity to induce variations in the function or its relationship. Nevertheless, students often face challenges in comprehending the nature and significance of the variable, primarily influenced by the notation employed (e.g., Dubinsky & Wilson, 2013; Olteanu & Olteanu, 2012). In this situation, teaching plays a crucial role in providing students with the opportunity to distinguish the meaning of the notation being used. The aim of this article is to provide insight into how this opportunity is presented in teaching that takes place in two upper secondary classrooms. The research question is: *How are different notations of the concept of a function presented in two upper secondary school classrooms, and what meaning do the students attribute to them?*

## THEORETICAL ASSUMPTIONS

The theoretical framework used in this study is the variation theory (Marton, 2015). The variation theory posits that learning entails experiencing something in a qualitatively novel and more profound manner. Among its core tenets, variation theory emphasizes the crucial role of students experiencing different aspects of the object of learning. An object of learning is a component of an educational situation that emerges through the interaction between the teacher and students, and it can be analytically divided into the enacted object of learning (what is observed in the classroom) and the lived object of learning (how students understand the object of learning) (Olteanu, 2016). The object of learning in this article is the role and the meaning of notations used to express a function.

A function is a mathematical object that can be represented in different ways, including arrow chains, tables, graphs, formulas, and phrases, each offering a different perspective on the same concept. In literature, two common approaches to interpreting and constructing functional relationships are frequently discussed: correspondence and covariation (e.g., Confrey & Smith, 1994). The correspondence approach is emphasized by the notation $y = f(x)$ and is also evident in specific teaching methods, such as input-output models. On the other hand, a covariation approach involves understanding how the dependent and independent variables change in relation to each other. To enable students to effectively comprehend the concept of a function, it is crucial for them to have a strong grasp of these approaches. This means that the communication between the teacher and students should focus on, among other things, helping students understand the meaning of the function's argument, the function's value, and the resulting outcomes.

To capture the students' attention regarding these aspects, the teacher may create situations that encourage exploration and combinations of them in various ways. This allows students to ascribe new meanings to each aspect and achieve a holistic interpretation of the concept of a function. However, if these aspects are not discerned by the students, they remain critical for students' learning. Aspects that have not yet been discerned and need to be discerned to enable learning are termed as critical aspects (e.g., Marton, 2015; Olteanu, 2018). Determining what constitutes critical aspects for a specific object of learning is always an empirical inquiry, and it is proposed that they can be pinpointed through tests, examination of textbooks and problems, or by conducting classroom observations (e.g., Marton, 2015; Olteanu, 2018). It is important to note that critical aspects should not be confused with difficulties, errors, or obstacles. Critical aspects lie between errors, difficulties, and obstacles and serve as conditions for progress and learning. It is also important to emphasize that critical aspects do not denote a lack of knowledge but rather the outcome of knowledge.

The study presented in this article centers on the same object of learning, and the data analysis explores how this object is developed within the classrooms.

## METHOD

The study was conducted in two classes from the Natural Science Program in upper secondary school. Both classes used the same textbook, and the study involved a total of 45 students (25 males, 20 females) who were 16 years old, along with two teachers named Anna and Maria. The teachers taught the same mathematics course. Data collection consisted of five steps: the students took a diagnostic test at the beginning of the course; the lessons were recorded on video; the students took two tests during the course and a diagnostic test after the course; eight students (four in each class) were selected for an individual session, including a post-test with tasks related to concepts they needed to further develop, and an interview; the teachers reviewed the video recordings, analysed the students' tests, and determined areas in which each student could improve their knowledge.

The results presented in this article are derived from the analysis of video-recorded lessons (12 lessons in each class). Since both teachers taught the same content using the same textbook, it was possible to identify and describe differences in their teaching approaches. The aim is not to draw comparisons between the two teachers who took part in the study, but rather to showcase and juxtapose two distinct approaches to delivering the same content. Additionally, by analysing the student's work during the teaching, it was possible to identify the meaning that students attribute to the content presented in the classroom. The analysis of collected materials aimed to identify different notations of the concept of a function presented in the classrooms, and what meaning students attribute to them.

## RESULTS

In the analysed video-recorded lessons, it was observed that the teaching in both classes focused on the notation $y, f(x)$ or $y = f(x)$ for functions. However, there were notable differences in how the graphical representation of functions and the corresponding notations were addressed.

In Maria's class, the representation $(y = f(x))$ was used without providing any explanation. She uses "$f(x) =$" as another way of writing "$y =$".

Maria: So instead of writing $y$, I write $f$ of $x$.

Maria primarily focused on the procedure for determining the value of the function, without delving into the meaning behind the notations. Maria introduces the concept of a function by emphasizing the function's argument and the function's value for specific arguments in various forms of representation, namely graphically, algebraically, or with the use of a table of value. At this point, Maria highlights that the notation $f(x)$ can be used instead of $y$. In this initial phase, Maria focuses on relating a function's argument to its value through the algebraic expression that constitutes the function's mapping rule.

In creating this relationship, she employs both $y$ and $f(x)$ as notations for the function's value. The difference between these notations is that the function's argument appears implicitly (in the notation with $y$) or explicitly (in the notation with $f(x)$), but Maria does not investigate this distinction in her discussion. Furthermore, in her presentation, Maria does not mention that x represents the function's argument and is an independent variable. The dual meaning of the notation "$y = f(x)$" poses problems for students as they work on their tasks. This is evident in, for example, the following conversation between Maria and Sune. The conversation is about the use of the following table to find a solution to the equation $f(x)$ = 29.

| $x$ | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| $f(x)$ | −10 | −7 | 4 | 29 | 74 | 145 |
| $g(x)$ | −6 | −7 | 6 | 33 | 74 | 129 |

[1] Sune: A solution to the equation?

[2] Maria: F of x is equal to 29 (reading the task in the book).

[3] Sune: Does that mean $3x$ is equal to $29y$?

[4] Maria: No, it doesn't mean that. F of $x$...

[5] Sune: The function of $x$...

[6] Maria: Yes, yes, but the function of $x$ is equal to $y$.

[7] Sune: Mm.

[8] Maria: So, this means that an equation solution where you get an $x$...

[9] Sune: So, three (shows what he wrote on his notepad)...

[10] Maria: Yes, you have, you have...

[11] Sune: Let me see, no, I made a mistake.

[12] Maria: To find the solution to the equation, you need to determine $x$. $x$ is the solution to the equation.

[13] Sune: Yes.

[14] Maria: In this case, $y$ is equal to 29...

[15] Sune: Mm.

[16] Maria: So, for which $x$ value is $y$ equal to 29?

[17] Sune: Three.

[18] Maria: Yes!

[19] Sune: Is that all I should write?

[20] Maria: Yes, you just need to write x is equal to three.

From the conversation, we can see that Sune struggles to understand the symbol $y = f(x)$ in this context [5]. This may be because Sune, on one hand, fails to recognize the relationship between different representations of a function, namely in the form of a value table including the notation $f(x)$, and on the other hand, he cannot differentiate the function's value at the point $x$ from the function's argument $x$. The dual meaning of the symbol $y = f(x)$ combined with the understanding of what is meant by the function's argument $x$ leads Sune to associate $f(x) = 29$ with the equation $3x = 29y$ [3]. Maria emphasizes that $f(x)$ is the same as $y$ and that solving an equation means finding the $x$-value that makes $y$ equal to 29 [4–13]. This enables Sune to connect his previous experience of solving an equation to the new context [17].

In contrast, Anna's class and the textbook utilized the notation $f(x)$ (explicit argument). Firstly, she utilizes a table of values through which a function's argument, mapping, and the function's value are presented. This is evident in the following transcript and what is written on the board:

Anna: I've taken $x$ and, what we say, mapped it to y, and then we can call it something (writes f above the arrow). This is a rule (points to $y = 2x + 3$) for how I calculate my $y$. I call it a function.

Using the provided value table, Anna emphasizes that the meaning of a function is a mapping, namely mapping x to y. Anna simultaneously presents the meaning of the notation f as a mapping and the relationship between this mapping, the function's argument, and the function's value for this argument. Anna also draws the graph of the function $y = 2x + 3$ on the board and uses the given function to calculate the function's value at point 2.

[38] Anna: I've taken $x$ and, as we say, mapped it onto $y$, and then I can call it something (writes $f$ above the arrow), this is a rule (points to $y = 2x + 3$) for how I calculate my $y$. I call it a function.

[…]

[59] Anna: If I want to calculate this, the value of this polynomial, the value of this expression for $x$ equals 2, it becomes a lot of words. So instead of now saying what the value of this expression is, what $y$ is when $x$ is equal to 2, I say it this way: $f$ of $x$ is equal to... the values of the function when $x$ is 2, that's what I usually say then, right.

The teacher writes on the board and speaks aloud: $f(x) = 2x + 3$ and

$f(2) = 2 \cdot 2 + 3 = 7$

From the graphical representation, it can be observed that Anna distinguishes the symbol $y$, which refers to the algebraic expression used to graphically represent a function, from the symbol $f(x)$, which refers to the algebraic representation of the function's mapping rule. In this way, Anna demonstrates that the function's argument can appear as explicit and implicit, yet the function maintains the same structure. However, Anna does not engage the students in a discussion about this difference. Additionally, she does not highlight that parentheses are used to indicate the function's argument, which differs from how parentheses are used in algebra. In this context, Anna highlights that $x$ is called the independent variable and y the dependent variable.

Anna consistently employed the notation $f(x)$ when engaging with different functions represented algebraically. This approach provided the students with an opportunity to observe variations in how the fundamental components of a function (argument, operation on the argument, and resulting output) interconnect and contribute to the overall concept of the function. Anna emphasizes that the symbol y can be present in an equation and used to identify the coordinates $(x, y)$ on the function graph. This introduces a variation in the function's argument, demonstrating that it can be represented explicitly as $f(x)$ to denote a function, or implicitly as $\mathrm{y} = \mathrm{x}^2 - 4\mathrm{x} - 5$, where its meaning can be interpreted as either a function or an equation. In Anna's class, the notations were presented separately, but with an emphasis on their shared meaning.

**The lived object of learning**

The results presented in this section relate to four tasks that were given in the students' tests. Tasks 1 and 2 involve calculations of the function value, while task 3 requires reading $f(2)$ from a drawn graph. The analysis reveals that both classes of students face challenges in using the concept of function both algebraically and graphically. However, it is noteworthy that students in Anna's class showed better performance compared to those in Maria's class.

Sixty percent of the students in Maria's class and eighty percent of the students in Anna's class could discern that the notation f(x) represents the value of the function f when given a value of x, along with the associated defining rule. A part of the students' presentations, 15% in Maria's class and 10% in Anna's class, indicate that the students distinguish the given functions as an operation without considering that $x$ represents the argument (the input) of the function, as the following example illustrates:

$f(x) = 12 + 4 = 16$ Svar: 16

$f(x) = 32 - (-5)^2$ Svar: 7

Another part of the students, 20% in Maria's class and 5% in Anna's class, did not discern the difference between the function argument and the function value and set up equations incorrectly, such as $4 = 12 + x$ and $-5 = 32 - x^2$.

The results indicate that a higher percentage of students in Anna's class were able to discern the function arguments and understand the relationship between the arguments and the function values, both algebraically and graphically.

## DISCUSSION AND CONCLUSION

Our approach examines how are different notations of the concept of a function presented in two upper secondary school classrooms and what meaning the students attribute to them. The results indicate that a critical aspect in both classes is the function's argument and the relation between explicit and implicit arguments in the notations $y$ and $f(x)$. This critical aspect is related to what Confrey and Smith (1994) referred to as the correspondence approach to functions. Despite classroom communication and textbook explanations, the argument of functions remains challenging for students to grasp, particularly in Maria's class compared to Anna's class. One possible reason for this disparity is Maria's use of implicit arguments without clear explanations or the presence of ambiguous relations between different components of the object of learning (function and equation). In contrast, Anna focuses on the critical aspect of the object of learning in her classroom. In Maria's class, students struggle to make sense of the signs, while in Anna's class, they can assign meaning to the signs more easily. It is alarming to observe that most of the students in upper secondary school do not discern the critical aspects related to the notation of functions.

Nonetheless, given that the concept of a function is fundamental in any mathematics course, comprehensive research has been carried out to address various issues related to learning this concept. Clearly, gaining insight into the students' learning process is intricately tied to understanding their challenges with the concept (e.g., Chen, 2023; Dogan-Dunlap, 2007; Thompson & Carlson, 2017). Therefore, this study underscores the importance of shifting the focus from students' difficulties to identifying the critical aspects related to the object of learning, as well as comprehending the meaning attributed to them by the students. This shift in perspective can enable teachers to better assist their students in grasping the concept of a function. Notations for functions should be unambiguous, as using the same notation for different objects hinders students' ability to discern the parts and relationships within a function. It is essential for teachers and textbook authors to clarify the purpose and significance of different notions, as this promotes comprehension and understanding.

### References

Chen, X. (2023). Study on Learning Difficulties of Function Concepts and Solutions. *International Journal of Education and Humanities*, *10*, 113–117. https://doi.org/10.54097/ijeh.v10i1.10935

Clement, L. L. (2001). What Do Students Really Know about Functions? *Mathematics Teacher*, *94*(9), 745–748.

Confrey, J. & Smith, E. (1994). Exponential functions, rates of change, and the multiplicative unit. *Educational Studies in Mathematics, 26*, 135–164.

Dogan-Dunlap, H. (2007). Reasoning with metaphors and constructing an understanding of the mathematical function concept. In *Proceedings of the 31st Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 209–216). PME.

Dubinsky, E., & Wilson, R. T. (2013). High School Students' Understanding of the Function Concept. *Journal of Mathematical Behavior*, *32*(1), 83–101. https://doi.org/10.1016/j.jmathb.2012.12.001

Marton, F. (2015). *Necessary Conditions of Learning*. Routledge

Minh, T.K. & Lagrange, JB. (2016). Connected functional working spaces: a framework for the teaching and learning of functions at upper secondary level. *ZDM Mathematics Education 48*, 793–807. https://doi-org.proxy.lnu.se/10.1007/s11858-016-0774-z

Musgrave, S., & Thompson, P. W. (2014). Function notation as idiom. *Paper presented at the Proceedings of the 38th Meeting of the International Group for the Psychology of Mathematics Education*, 2. PME.

Olteanu, C. (2018). Designing examples to create variation patterns in teaching. International *Journal of Mathematical Education in Science and Technology, 49*(8), 1219–1234. https://doi.org/10.1080/0020739X.2018.1447151

Olteanu, L. (2016). *Framgångsrik kommunikation i matematikklassrummet*. Diss. Linnaeus University Press.

Olteanu, C., & Olteanu, L. (2012). Equations, Functions, Critical Aspects and Mathematical Communication. *International Education Studies*, *5*(5), 69–78.

Sajka, M. (2003). A secondary school student's understanding of the concept of function—A case study. *Educational Studies in Mathematics*, *53*(3), 229–254. https://doi.org/10.1023/A:1026033415747

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.

Thompson, P. W., & Milner, F. A. (2019). Teachers' meanings for function and function notation in South Korea and the United States. In H.-G. Weigand, W. McCallum, M. Menghini, M. Neubrand & G. Schubring (Eds.), *The Legacy of Felix Klein* (pp. 55-66). Springer.

Trujillo, M., Atarés, L., Canet, M. J., & Pérez-Pascual, M. A. (2023). Learning Difficulties with the Concept of Function in Maths: A Literature Review. *Education Sciences*, *13*(5), Article 5. https://doi.org/10.3390/educsci13050495

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

THE UNIVERSITY OF
AUCKLAND
Te Whare Wānanga o Tāmaki Makaurau
NEW ZEALAND

AUT
TE WĀNANGA ARONUI
O TĀMAKI MAKAU RAU