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 4
Research Reports
(P – Z)

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 4

Research Reports
(P – Z)

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. 4). 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: 978-1-0670278-4
ISSN: 0771-100X

Logo designed by Jason Lamontanaro

TABLE OF CONTENTS – VOLUME 4

REASONING QUANTITATIVELY ABOUT THE LUNAR PHASES

Nicole Panorkou and Amanda Provost

Montclair State University

We discuss how students' reason quantitatively as they explored the lunar phases in an interactive computer simulation we designed to integrate math and science learning. We illustrate different forms of students' quantitative operations as they reason about the moon in its orbit.

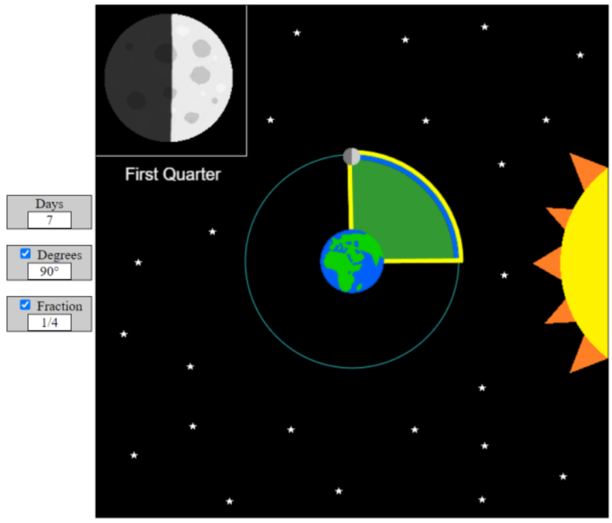
As the importance of integrated education continues to grow, more research is needed to determine how to strengthen the reciprocal relationship between mathematics and science by providing meaningful learning opportunities where the mathematics support the science learning, and the science learning supports the mathematics learning (Fitzallen, 2015). In our prior work, we have shown that students can express sophisticated forms of mathematical reasoning while engaging with the activities and dynamic simulations we have designed to integrate mathematics content into earth and environmental middle school topics (e.g., Panorkou & Germia, 2020). In line with these efforts, recently we presented a preliminary analysis on how one pair of students have reasoned covariationally and multiplicatively about the lunar phases (Provost & Panorkou, in press). This paper expands this work by examining the reasoning of three more students and comparing this reasoning to our previous preliminary analysis.

BACKGROUND

The scientific phenomenon of the lunar phases is a prominent curriculum topic in middle school in the United States (NGSS Lead States, 2013). This is often a challenging topic for students as they usually have alternative conceptions for the cause of the lunar phases such as believing that some object blocks part of the moon (e.g., Wilhelm et al., 2022). To address this difficulty, our research group designed the Moon Pie simulation (Figure 1a) to model the relationship between the moon's revolution around the earth and its different phases. The user can drag the moon around in its orbital path, observing the resulting changes in its phase as displayed in the picture-in-picture view of the moon as seen from earth. The simulation displays a readout of the number of days elapsed during the moon's orbit using a 28-day approximation of a lunar month. (Note that it takes the moon approximately 29.5 days to travel around the earth.) To support students to make a connection between the phases and the position of the moon in relation to a full orbit, the simulation presents the moon's travel measured in days, degrees, and fractions of a full orbit. For example, Figure 1a shows the First Quarter moon on Day 7 of this 28-day month. Students can toggle overlays displaying measurements of the moon's progress through its orbit in both degrees (blue overlay) and fractions (yellow overlay) to see how these measurements are related to each other and to the number of days elapsed. When both overlays are toggled on, they combine to form a green overlay. Figure 1b shows the quantities of time of travel (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. 4, pp. 1–8). PME.

days), wedge of the orbit (as a fraction), and arc of the orbit (in degrees) that are associated with each lunar phase. We also designed probing questions to support students' reasoning of mathematical relationships as they explore the simulation, such as: *What do you notice?*, *What is changing? How is it changing?*, *What fraction of the orbit is it from [Lunar Phase] to the [Lunar Phase]?*, and *How many degrees of the orbit does it take for the moon to get from [Lunar Phase] to [Lunar Phase]?*



Phase	Day	Fraction	Degree
New Moon	0	0	0°
Waxing Crescent (midpoint)	3.5	$\frac{1}{8}$	45°
First Quarter	7	$\frac{1}{4}$	90°
Waxing Gibbous (midpoint)	10.5	$\frac{3}{8}$	135°
Full Moon	14	$\frac{1}{2}$	180°
Waning Gibbous (midpoint)	17.5	$\frac{5}{8}$	225°
Third Quarter	21	$\frac{3}{4}$	270°
Waning Crescent (midpoint)	24.5	$\frac{7}{8}$	315°
New Moon	28	1	360°

Figure 1: The Moon Pie simulation (left). The measures in each Lunar Phase (right)

In our preliminary analysis (Provost & Panorkou, in press), we examined the quantitative reasoning of one pair of students as they engaged with the Moon Pie simulation. We refer to a quantity (Thompson, 1994) as a conceived attribute of an object or phenomenon that is measurable. For example, students reasoned about the position of the moon in its orbit or distance travelled in terms of the quantities of the fraction of the circular orbit traced or the degrees of the arc generated. Students reasoned about the variation in a single quantity and coordinated the simultaneous variation in multiple quantities, engaging in covariational reasoning (Thompson & Carlson, 2017). For instance, students reasoned that as the moon moves in its orbit, the arc of the moon's orbit measured in degrees increases, and the number of days of travel also increases. Students also used co-splitting, a form of covariation in which any multiplicative change in one quantity is coordinated with the same multiplicative change in the other quantity (Corley et al., 2012). For example, they reasoned that when the arc of the moon's orbit triples, the fraction of the orbit also triples. Overall, the Provost and Panorkou (in press) study illustrated some forms of reasoning that are possible when students engage with the Moon Pie simulation and our questioning. This paper presents our further analysis with more students that provided additional evidence to support these forms and shed light into some novel forms of reasoning not reported on before. Specifically, we examined: *In what ways do middle school students reason quantitatively about lunar phases when utilizing our simulation design, and how does this compare to our preliminary findings?*

METHODS AND ANALYTICAL FRAMEWORK

We analysed the reasoning of three more students from the same whole-class design experiment (DE) (Cobb et al., 2003) as the Provost and Panorkou study, which was conducted in a sixth grade (11-12 year old) classroom in northeastern U.S.A. Students spent two 40-minute sessions working with the Moon Pie simulation. These sessions were conducted over Google Meet due to COVID-19 and were video-recorded.

This paper reports on how the reasoning of this new group of students, Gaelyn, Jami, and Ghina, compares to our previous analysis. In addition to covariational reasoning, we were particularly interested in the forms of quantitative operations (Thompson, 1994) that students used. For example, students may use repeated addition (Fischbein et al. 1985), such as reasoning about the distance from the New Moon to the Full Moon in terms of 45° + 45° + 45° + 45° = 180°. Students may also reason multiplicatively using the coordinated measurement approach (Izsak & Beckmann, 2019) of $N \times M = P$, where N is the number of degrees, days, or a fraction of orbit in one group, M is the number of groups the distance from two phases makes, and P is the total distance from two phases. For instance, reasoning about the distance from the New Moon to the Full Moon in terms of 45° per group × 4 groups = 180°. We also examined whether they exhibited one-dimensional forms of reasoning (single quantity) or if they constructed any multiplicative objects (Thompson & Carlson, 2017), which entails the coupling of two quantities from the Moon Pie simulation. As Thompson and Carlson (2017) describe this as: “A person forms a multiplicative object from two quantities when she mentally unites their attributes to make a new conceptual object that is, simultaneously, one and the other” (p. 433).

FINDINGS

We first report on the quantities students constructed as they used the Moon Pie simulation and then discuss how they reasoned about single or multiple quantities, what we refer to as one- or two-dimensional forms of reasoning.

Constructing quantities

Students were asked to describe what is changing in the simulation and how. The students constructed the position of the moon as the controllable quantity, and discussed the resulting changes in time in orbit, wedge of the orbit, and arc of the orbit. For instance, Jami stated, “when you move the moon [moved it to Full Moon], the fraction changes at a certain point. For example, this is one-half, and then also the degrees changes.” Gaelyn described, “if we change the position of the moon, that will change the name of lunar phase and it will also change the number of days and the visible part of the moon, the degrees of moon’s orbit, and the fraction of moon’s orbit.”

One-dimensional forms of reasoning

Students were then asked a series of questions that probed them to identify the fraction travelled from New Moon to First Quarter to Full Moon to Third Quarter to New Moon. They used additive reasoning to identify that the fraction of the orbit’s wedge travelled between each of these phases would be one-fourth. For instance, from the First Quarter to Full Moon, Gaelyn explained that,

> "it will be one-fourth because the Full Moon is one-half of the whole orbit, of all the phases, and one quarter is one-fourth so subtract one-quarter from one-half that will be another quarter."

Here Gaelyn had identified the fraction of the orbit's wedge for the Full Moon and First Quarter moon phase positions and used subtraction to find their difference. (Note her use of "one quarter" instead of First Quarter.) Similarly, when asked about the fraction of the orbit travelled from First Quarter to Third Quarter (Figure 2), she stated, "Well, what I thought is one-fourth to three-fourths and the difference between that it's two-fourths, and two-fourths as a fraction is equivalent to one-half." Ghina also used subtraction of fractions reasoning that, "three-quarters minus one-quarter is two-quarters, which is one-half." In a similar manner, when asked to identify the degrees travelled from the same two phases, Gaelyn and Ghina subtracted the difference in degrees between the two positions. As Ghina explained, "it's going to be 180 because when you're at First Quarter it's already 90° and Third Quarter is 270°, and the difference between 270 and 90 is 180." Likewise, Gaelyn argued, "Well, it would be from 90° to 270° and the difference between that would be 180°."

The third student, Jami, constructed a quarter of the orbit as a unit which she iterated using repeated addition stating, "So First Quarter and then we have to get to Third Quarter, [moved the moon to the Full Moon position] that's one quarter and then [moved the moon to the Third Quarter position] that's two quarters which is one-half." Likewise, to find the degrees travelled, Jami iterated by 90° explaining,

> "Also, I think it's 180 because when you're at First Quarter [moved moon to First Quarter] and then it's 90° to get to Full Moon [moved moon to Full Moon] which is the next lunar phase, and then 90° to get to Third Quarter [moved moon to Third Quarter] which is 90 plus 90 is 180. So, it's when you're traveling 180° to get to Third Quarter."

Jami's statements show that she recognizes a quarter wedge to be equivalent to one-fourth of the orbit and 90°, and uses this as a unit of measure to find the distance travelled between phases (Figure 2a):

> "Full Moon to the next First Quarter, I said it was three-fourths [moved the moon to Full Moon position] because if you go, so from Full moon to [moved the moon to Third Quarter position] Third Quarter, that would be a one-fourth, then two-fourths [moved the moon to New Moon position] from Third Quarter to New Moon. And then three-fourths would be [moved the moon to First Quarter position] from New Moon to First Quarter."

On the contrary, Ghina and Gaelyn's reasoning showed that they constructed one-half of the orbit or 180° as a composite unit that they used whenever they were asked to find a larger distance. For example, for the same distance, Gaelyn stated:

> "I know that the Full Moon is at one-half and to finish the next cycle, it would be another one-half so, and then to finish up to another First Quarter would be one-fourth, I would add one-half and one-fourth together and get three-fourths."

Likewise, for finding the degrees, Ghina explained,

> "I was going to say 270°. Because when you're like the Full Moon, and then you head to New Moon that's 180°, I'm pretty sure. And then, and then you go travel another 90° to First Quarter, and 90 plus 180 is 270."

The three students also reasoned multiplicatively using coordinated measurement. For example, when asked about the total number of degrees travelled to complete two full orbits, Gaelyn and Ghina showed that they even constructed 360° or a full orbit as a composite unit (a group) that they multiplied by two. As Gaelyn explained, "Because a circle is 360° and that's when it orbits twice, you have to multiply 360 by two and that would be 720." Ghina gave a similar explanation, "Since we know that, to get it done once 360°, so we just multiply by two to get 720° since we want two times."

Similarly, Jami used a coordinated measurement approach to explain that her unit (group) of 90° can be multiplied by four to get the total 360° of the orbit. When asked how the math helps with the science in the simulation (Figure 2b), she explained,

> "The math helps us with the science because if you know, a 90°, it's like 360° and 90° angles, you'll know that 90°, well 90° times four is 360. And then there's four main points [moved the moon to each phase as she mentioned it]: the New Moon, the First Quarter, the Full Moon, the Third Quarter, and the next New Moon. So, [moved the moon to each phase she mentioned] if you know then the First Quarter is 90°, second quarter is Full Moon, the Third Quarter is another 90°, and the next New Moon is 360° which is plus another 90."

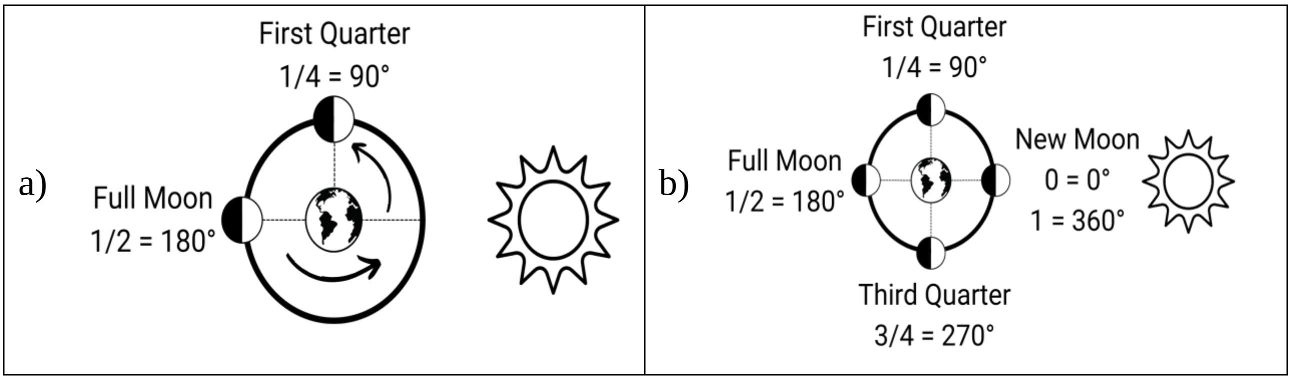


Figure 2: a) Travel from Full Moon to next First Quarter; b) The 'four main points'.

While students were using either fractions or degrees in their explanations based on the question asked, their reasoning showed traces of coupling the quantities of fraction of the orbit and the degrees of the orbit. For instance, in the statement above, Jami's use of "second quarter" to refer to Full Moon shows that her construction of the unit entails both quantities. Also recall Gaelyn's use of "one quarter" instead of First Quarter at the beginning of this section. The next section describes their coupling of quantities as multiplicative objects in more depth.

Two-dimensional forms of reasoning

As they explored the simulation, students first coupled the quantities of the wedge of the orbit and the time in orbit. For instance, when discussing the quarter phases (First Quarter, Full Moon, Third Quarter, New Moon), Jami stated,

> “they’re using instead of like the 28 days, they are using fourths to simplify it. … Because the moon phases are broken up into quarters so that’s why the first quarter gets its name, and so using for example 7/28 you would simplify it down to one-fourth.”

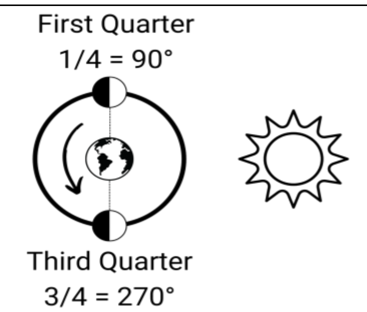
Jami elaborated on this relationship later by using equivalent fractions,

> “And you can also use them [fraction of the orbit] for 28 days, you can make the denominator 28 instead of four, and you can make the numerator the correct day. So, for example, First Quarter would be seven-twenty eighths.”

We interpret her reasoning to show that she recognizes that there is a fractional measure of the wedge for every measure of the time in orbit in days, illustrating the simultaneous coupling of the two quantities. Likewise, when Gaelyn was discussing the moon travelling from First Quarter to Third Quarter she stated (Figure 3),

> “because I know from First Quarter to Third Quarter it’s another way of just writing a line vertically down the middle and that would be one-half and the total degrees of a circle is 360 and a half of that would be 180.”

Both students’ statements show that they performed the same operation on both quantities simultaneously. For instance, Gaelyn coordinated that splitting a circle in half entails co-splitting the arc of a full orbit in half.



Co-splitting

Fraction	Degree
1	360°
1/2	180°

× 1/2 × 1/2

Figure 3: Gaelyn’s reasoning about travel from First Quarter to Third Quarter moon.

We noticed that students’ coupling of the quantities of wedge of the orbit to the arc of the orbit led to a construction of two-dimensional units of measure. For instance, when asked how the fraction of the orbit is related to the number of degrees travelled, Gaelyn responded, “I found out that 90° is how much every one-fourth is. That’s the pattern.” We interpret Gaelyn’s statement to show the construction of a unit that includes two quantities (one-fourth = 90°). Students then used this two-dimensional unit to reason about the moon travelling around the orbit. For example, Jami stated:

> “Another pattern is that. So, every 90° is one-fourth. So, for example, 90° is the First Quarter, another 90°, which is 180°. When you add it, it’s the Full Moon. And 180 is half of 360. And that’s half of, two-fourths is half. That’s like half the moon’s orbit. And then you add another 90°, which is now 270°, and it’s the Third Quarter. And it completed three-fourths of its orbit. And then another 90° and it’s 360° and it has completed its orbit.”

Here Jami was using repeated addition to iterate the two-dimensional unit one-fourth = 90° around the orbit (Figure 4). Note that as she was iterating by fourths around the orbit, she also used co-splitting to show evidence of constructing the relationship between one orbit = 360° and one-half orbit = 180°. We interpret her statement that "two-fourths is half" to illustrate the beginning of a construction of a composite unit of one-half orbit = 180° that consists of two units of one-fourth orbit = 90°.

	Fraction	Degree	
	1/4	90°	
+ 1/4	2/4 = 1/2	180°	+ 90°
+ 1/4	3/4	270°	+ 90°
+ 1/4	1	360°	+ 90°

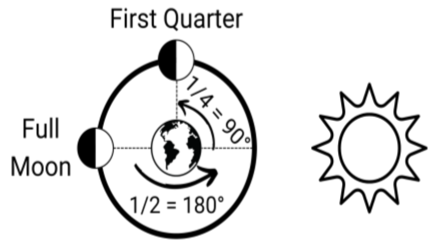
	Fraction	Degree	
	1	360°	
× 1/2	1/2	180°	× 1/2

Figure 4: Jami's two-dimensional repeated addition (left) and co-splitting (right)

Ghina also showed evidence of a construction of a composite unit when she was asked about the fraction of the orbit travelled from the Full Moon to the next First Quarter (Figure 5a) and stated,

> "So, when you're at the Full Moon and you travel 180°, that's one-half. And then you travel another 90°, that's one-fourth. And one-half is also two-fourths. So, you can just do two-fourths plus one-fourth to just three-fourths, which is also 270°."

a)



b)

	Fraction	Degree	
	1	360°	
÷ 4	1/4	90°	÷ 4
× 1/2	1/8	45°	× 1/2

Figure 5: a) Ghina's reasoning about the Full Moon to First Quarter distance and b) Gaelyn's reasoning using co-splitting to find eighths.

Similarly, Gaelyn used co-splitting to split both quantities even further and construct an even smaller two-dimensional unit. For instance, she noticed (Figure 5b), "Everyone one-eighth fraction is 45°. …Well, because one-fourth is 90°. And half of one-fourth is one-eighth, and a half of 90° is 45°."

CONCLUDING REMARKS

This paper discussed three students' quantitative reasoning as they explored the lunar phases and travel of the moon. Similar to our preliminary analysis, students reasoned about the simultaneous variation in multiple quantities (Thompson & Carlson, 2017) and used co-splitting (Corley et al., 2012) to operate on two quantities.

In addition to these forms of reasoning, the current study provides more depth on the nature of the quantitative operations that students used. The design of the simulation illustrating both the wedge as a fraction of the circular orbit and the arc of the orbit in degrees showed to support students' coupling of these two quantities as a multiplicative object. This coupling led to the construction of two-dimensional units, and two-dimensional composite units, that they used in quantitative operations to discuss how the moon travelled. Our findings showed evidence of a two-dimensional form of repeated addition, where students iterated a two-dimensional unit. We have also seen traces of a novel form of the coordinated measurement approach in which one "group" is defined as a two-dimensional unit. In the future, we plan to examine whether other students have used similar operations and study these new forms of reasoning further.

Acknowledgments

This research was supported by the National Science Foundation (#1742125). The data presented, statements made, and views expressed are solely the authors' responsibility.

References

Cobb, P., Confrey, J., DiSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher, 32*(1), 9–13.

Corley, A., Confrey, J., & Nguyen, K. (2012). The co-splitting construct: Student strategies and the relationship to equipartitioning and ratio [Paper presentation]. *American Education Research Association Annual Meeting*, Vancouver, Canada.

Fischbein, E., Deri, M., Nello, M. S., & Marino, M. S. (1985). The role of implicit models in solving verbal problems in multiplication and division. *Journal for Research in Mathematics Education, 16*(1), 3–17.

Fitzallen, N. (2015). STEM education: What does mathematics have to offer? In M. Marshman, V. Geiger, & A. Bennison (Eds.). *Mathematics education in the margins (Proceedings of the 38th annual conference of the Mathematics Education Research Group of Australasia)* (pp. 237–244). Sunshine Coast: MERGA.

Izsak, A., & Beckmann, S. (2019). Developing a coherent approach to multiplication and measurement. *Educational Studies in Mathematics, 101*(1), 83–103

NGSS Lead States. (2013). *Next Generation Science Standards: For States, By States*. The National Academies Press.

Panorkou, N., & Germia, E. F. (2021). Integrating math and science content through covariational reasoning: the case of gravity. *Mathematical Thinking and Learning, 23*(4), 318-343.

Provost, A., & Panorkou, N. (in press). Exploring lunar phases with the Moon Pie simulation. *Mathematics Teacher: Learning and Teaching PK-12*.

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. 181–234). 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.

Wilhelm, J., Cole, M., Driessen, E., Ringl, S., Hightower, A., Gonzalez-Napoleoni, J., & Jones, J. (2022). Grade level influence in middle school students' spatial-scientific understandings of lunar phases. *School Science and Mathematics, 122*(3), 128-141.

LEARNING ROUTES FOR ALGEBRAIC THINKING IN PRESCHOOL

Elena Polotskaia*, Nathalie Silvia Anwandter Cuellar*, Annie Savard**, and Virginie Robert***

*Université du Québec en Outaouais, **McGill University, ***ULaval

In this theoretical essay, we use observations from our ongoing research with preschool children to question the theoretical frameworks available for studying the developmental trajectories of algebraic thinking in young children. We critically analyse two approaches. The theoretical approach employed by Early Algebra presumes that elementary students develop algebraic thinking by using some knowledge of numbers and arithmetic operations. The theoretical approach employed by Davydov, and his followers presumes that the most general ideas of algebraic thinking are prerequisites for the study of numbers and operations. How do these approaches interplay to allow for an interpretation of what we observe in preschool?

OUR OBSERVATIONS

For the ongoing research project with 4-5-year-old children (preschool) we created learning activities in a game-like format to elicit young children's algebraic thinking. We created four games within the context of weight. In the *first game*, children try to feel with their hands the weight of different objects and discuss this and other characteristics of those objects. For example, two apparently similar bottles of yogurt weigh differently; one is heavier, and the other is lighter. The *second game* allows the students to freely use a two-plates scale to compare objects and groups of objects by their weights. The *third game* uses specially created materials: plastic blocks of four different colours identified by animal stickers—mouse, cat, dog, and pig. The weights of the blocks correspond to each other as follow: cat = 2 mice; dog = 3 mice; pig = 4 mice. A playing card presents several animals indicating the blocks to use. A player needs to distribute indicated blocks on a two-plates scale to obtain equilibrium. The *fourth game* employs cards presenting a two-plates scale with some animals on it in equilibrium. One animal on the scale is hidden. The player is required to identify the missing animal: predict and then verify this prediction on the scale. The games three and four are inspired by the works of Papadopoulos and Patsialia (2018). The following excerpt present the interaction between the teacher and two children playing the fourth game. It is the teacher’s turn to play. He must find one animal to equilibrate a mouse and a cat (see Figure 1).

Teacher: So here, I have one mouse and one cat (placing corresponding blocks near the card). I need something on the other side. One mouse, one cat. What is on the other side?

Clara: (Takes a dog from the pile of blocs).

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. 4, pp. 9–16). PME.

Teacher: One mouse, one cat.

Figure 1: Visual equation with Mouse and Cat on the left side of the scale and a hidden animal on the right side.

Clara: A dog.

Teacher: A dog on the other side? We will try. (Placing blocs on the scale according to the Clara's suggestion. The scale is in equilibrium; the dog weighs the same as a cat and a mouse together.) How did you know?

Clara: I knew.

Teacher: You knew? Please, explain, how did you know?

Clara: Because mice are lightweight.

Teacher: Mice are lighter, so you need something heavy.

Clara: But not too heavy!

Teacher: So, a pig would be too heavy? If I replace the dog by a pig? (Replaces a dog by a pig on the scale. The scale shows that the pig is heavier than a cat and a mouse together.) Yes, you are right. So, a bit heavier, but not too much. Wonderful!

METHOD

In our literature review (Polotskaia et al., 2019), we identified two major theoretical approaches to the introduction of algebraic ideas in elementary school: *Early Algebra* and *Developmental Instruction*. In the following sections, we present each approach, and we attempt to interpret our observations within these frameworks. Based on this analysis we propose a new framework for interpreting algebraic thought in preschool children.

FROM THE EARLY ALGEBRA POINT OF VIEW

The Early Algebra movement emerged out of the necessity to address one of the fundamental problems in mathematics education, namely the arithmetic–algebra gap. By the end of the 20th century, it became evident that a significant number of secondary students faced challenges in learning algebra. For instance, many students interpreted '=' operationally as 'calculate the result' rather than relationally as 'left and right expressions are equivalent' (see more examples in Kieran, 2018). Solutions to the arithmetic–algebra gap were proposed to facilitate a seamless transition from arithmetic to algebra by introducing activities that elicit algebraic reasoning at an earlier stage in elementary school (e.g., Kaput, 2008; Kieran, 2018).

Efforts were made to conceptualize the development of algebraic reasoning in students during the early years of schooling. Blanton and her colleagues (2018a, 2018b) use a definition of algebraic thinking formulated by Kaput (2008) to establish a theoretical framework for investigating the developmental routes of algebraic thinking in young students. According to Kaput (2008), algebraic thinking involves generalizing and communicating these generalizations in symbolic form, as well as using symbolic representations to apply the corresponding rules of the semiotic system. For young children, natural language and some culturally accepted visual representations are included as semiotic representations. Blanton et al. conclude:

> We derive four essential practices from Kaput's (2008) core aspects that define our early algebra conceptual framework: generalizing, representing, justifying, and reasoning with mathematical structure and relationships... (Blanton et al., 2018b, p.30).

In the experiments, the learning activities employed to elicit algebraic thinking were based on numerical expressions and their visual representations, where a number is depicted as a set of dots. Blanton et al. (2018a) propose the following classification of observable types of thinking about ' = ' in kindergarten students.

1. Operational: For example, interpreting an expression like 2 + 3 = 5 as combining or totalizing two numbers.
2. Emergent Relational: For example, using a nonstandard equation (5 = 3 + 2) while still interpreting ' = ' as totalizing/combining.
3. Relational: For example, using ' = ' to express the equivalence of values, evaluating whether the expression employing ' = ' is true or false, and solving equations for missing values.

We tried to use this approach to interpret Clara's reasoning. In our experiment, the context is not numerical; students do not need to count, as not more than four similar objects are used at a time, and the weight of each object is not expressed numerically. Equivalence is introduced as 'scale in equilibrium' or 'scale in an even position.' The students did not express their understanding in writing or drawing; thus, ' = ' was not used. However, some kids spontaneously used their hands to show the position of the scale, or they used words like 'similar' and 'even,' etc.

Considering all these conditions, how can we interpret Clara's solution to the missing animal problem? Certainly, Clara is solving an equation and not 'adding' the cat and mouse to obtain the dog. The explanation of her choice of animal is not based on counting (e.g., dog = 3 x mice, and cat = 2 x mice) or on memorized knowledge (e.g., dog = cat + mouse). Instead, she uses approximate qualitative arguments. Her thinking probably goes as follows: she needs an animal a bit heavier than a cat but not too heavy. She explains that *mice are lightweight*, so the animal should not be *too heavy*. We can suggest that Clara's reasoning is strongly guided by the understanding of the situation as equilibrium or equivalence. We can conclude that Clara's thinking can be classified as relational. This reasoning is approximate and not numerical, and the ' = ' is not used.

We found that it was useful but difficult to interpret Clara's reasoning within a theoretical frame emerged from numerical contexts.

FROM THE DEVELOPMENTAL INSTRUCTION POINT OF VIEW

The Developmental Instruction approach proposed by Davydov (1982, 2008) is rooted in Vygotskian views of learning as a culturally formed, adult-child joint activity. This perspective implies that what students learn depends on the learning experiences offered by the teacher and the school tradition. If the traditional teaching of mathematics begins with counting and number operations, students have no other choice but to base their algebraic thinking on their knowledge of numbers and operations. Davydov suggests a fundamental shift in the school mathematical tradition by prioritizing learning about basic quantitative relationships—equality and equivalence—before the formal study of numbers and operations.

Furthermore, Davydov (2008) introduces the notion of theoretical thinking, characterized by understanding a structure or relationship between quantities, expressing, and communicating this relationship, and using it to logically deduce new information, such as solving a problem or finding a missing element in the relationship. Considering this definition, Clara's reasoning appears to be theoretical. The problem of the missing animal is represented by a schema or visual equation, and Clara mentally solves it without resorting to counting or manipulation. Clara understands that the weights on each side of the scale should be equivalent. To find a solution, she conjectures that to counterbalance a cat and a mouse, the other animal should be *heavier than a cat but not too much*, as *mice are lightweight*. This reasoning employs relationships between quantities to logically deduce a correct solution.

However, the relationships Clara employs are not the complete relationships as defined by Davydov as a basis for mathematical knowledge development. In his view, a relationship is a structure of three elements, where each element can be found or constructed when the other two are known. For example, the numbers 2, 3, and 5 are in an additive relationship because $2 = 5 - 3$, $3 = 5 - 2$, and $5 = 2 + 3$. In the case of the missing animal problem, the weight of the dog is equal to the sum of the cat's and the mouse's weights—an additive relationship according to Davydov. However, Clara does not explicitly know this relationship. Instead, she relies on approximation using partial relationships: the mouse is light (in relation to other animals), and the dog is not too heavy. These partial relationships do not constitute relationships of three elements.

Furthermore, Clara does not represent her understanding of the relationships with symbols or models. Oral communication of her reasoning is not entirely complete and clear. Evidently, Clara is at the beginning stages of developing the mathematical language (semiotic system) necessary for the situation.

We found it useful to interpret Clara's reasoning by employing the notion of theoretical thinking. Nevertheless, some theoretical elements are missing to adequately capture the mathematical essence of students' thinking.

DISCUSSION

We employed two different theoretical approaches with an example of a 5-year-old student's thinking while solving a missing animal problem. Each of these approaches provides a comprehensible structure to identify some characteristics of the student's thinking in association with algebraic or theoretical thinking.

Our analysis of Clara's reasoning shows that at the heart of both approaches is a deep understanding of a structure, equivalence, or relationship, working with the structure or structures (their meaning) to draw logical conclusions. We suggest that the definition of *theoretical thinking* seems to be very close to *algebraic thinking* employed in Early Algebra. In both cases, a deep understanding is a foundation, the use of symbolic communication is an essential part, as well as logically deducing a solution or justifying it. Yet, two major questions remain:

1. What are the contexts available to 4-5-year-old children in which structures or relationships can be studied and understood?
2. What semiotic systems, communication tools can young children employ to express their thinking and to move this thinking forward?

The experimenters in the Early Algebra approach use small numbers to discuss algebraic concepts with students (e.g. numerical expressions, solving word problems with numerical data, function machines, etc.). Thus, we can suggest that numbers and operations on (small) numbers are the main context used in Early Algebra.

The Developmental Instruction approach does not require number knowledge from students at the beginning. Even though some children may have some of this knowledge already developed, this knowledge is not really used to establish or employ simple relationships between continuous quantities. In multiple experiments (e.g., Mellone et al., 2018; Eriksson and Eriksson, 2020; Davydov, 1982), it was shown that young children can work with continuous quantities (volume, length, etc.) without counting or numerical expressions. Children can understand partial relationships between quantities (<, >, =), express those relationships using letter notations, produce logical conclusions, and justify their solutions. In this approach, contexts are physical objects and their quantitative characteristics, not expressed in numerical form. There are many such contexts accessible for young learners: lengthy objects (length), containers (volume), objects of different weight or area. Those contexts make some algebraic ideas (e.g. equivalence, using letters, evaluating expressions as true/false, etc.) available for kids to study earlier. In our experiment, we observed children who couldn't count to 10 but were able to solve some problems in weight games. So,if non-numerical contexts are used to study relationships and structures, the ways of communication and representation (modelling) should also be non-numerical. Researchers in Early Algebra (e.g. Blanton, 2018; Radford, 2011; Boily et al., 2020) showed that 4-8-year-old children can express their reasoning about relations by using special words and gestures. For example, to express equality, a child can say 'the same' or 'balance' (Blanton, 2018); to express the growth of a non-numerical sequence, a child can say 'going up' and produce a corresponding gesture (Boily et al., 2020); a child can use gestures to attract attention to general aspects in several structures (images) (Radford, 2011). In our experiment, the teacher introduced a special hand-gesture to express the position of equilibrium of the scale, and the children easily adopted this gesture and used some others for the same purpose. In the experiments by Davydov and his followers, letter notation and schematic representations were employed from the beginning and helped students to express their understanding (for more discussion about the use of letters in elementary grades, please see Tremblay et al., 2021). However, those communication tools are different in the level of abstraction they introduce, and some of them are unknown for children in preschool.

Taking into consideration that students' mathematical thinking (arguments) can employ complete or partial understanding of structures or relationships, and that students can express their thinking by using semiotic systems (communication tools) of different levels of abstraction, we propose the following bi-dimensional frame to analyse preschool children's algebraic thinking in non-numerical contexts (e.g., weight games). The first dimension reflects the quality of relationships a student uses to argue: partial relationships (bigger, smaller, identical) or complete relationships in Davydov's sense. The second dimension reflects the level of abstraction of the communication tools (semiotic systems) the student employs: direct manipulation without comments, use of natural language and/or gestures to express their thinking, argumentation using a culturally shared formal semiotic system (other than natural language). Table 1 presents what can be observable in the case of weight games.

Relationships used to construct an argument / Communication tools	**Partial relationships bigger, smaller, equal (identical objects)**	**Relationships of type *equation* (equality of groups of non-identical objects)**
Direct manipulation with objects (without comments)	Try and error strategy seems to be organized; Student directly chooses a correct object.	Impossible to judge without explicit arguments.
Argumentation in natural language, gesture	Student names or/and shows by gestures essential partial relationships.	Student names or/and shows by gestures essential full relationships
Argumentation in culturally shared semiotic system	Student represents partial relationships symbolically or schematically and uses these representations to advance her reasoning	Student represents full relationships symbolically or schematically and uses these representations to advance her reasoning

Table 1: Bi-dimensional model to interpret young students algebraic thinking in the context of weight games.

Direct manipulation with objects – partial relationships: In the missing animal game, Clara takes a dog from the pile of blocks (which is correct to counterbalance a cat and a mouse). Without oral arguments, it is impossible to judge what type of relationships is used. Another strategy would be to try animals one by one in order of their weight. This strategy suggests the use of partial relationships.

Argumentation in natural language, gesture – partial relationships: Clara explains simple relations: mice are lightweight, not too heavy, to argue her solution. Students can also try with their hands which animal is heavier or lighter.

Argumentation in natural language, gesture – full relationships: In a different case, a student can argue, for example, that a dog weighs the same as a cat and a mouse, so we can use a dog instead of a cat and a mouse, and the equilibrium will not change.

Argumentation in culturally shared semiotic system – partial relationships: In another case, a student could perform the following as an argument: D > C, so D + M > C + M, meaning that if one adds a mouse on each side of the scale, the scale will remain in the same position.

Argumentation in culturally shared semiotic system – full relationships: In a different case, a student can solve the following problem (see Figure 2) by using a schema and show that the missing animal should be a cat.

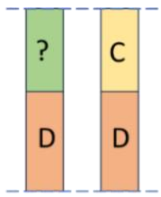


Figure 2: Visual equation with Dog and Cat solved by a schematic representation.

The proposed theoretical frame describes types of thinking 4-5-year-old children might use in terms of observable elements. These elements extend the definition of algebraic thinking employed in Early Algebra and Developmental Instruction approaches to include the understanding of partial relationships and the use of communication tools available for young kids but not yet developed to represent a fully formed semiotic system. We propose that the development of algebraic thinking in young children can go in two directions: from partial to holistic understanding of quantitative relationships and employing more and more formal (abstract) and complete semiotic systems.

Considering this bi-dimensional frame, Clara's reasoning can be classified as *argumentation in natural language, gesture by using partial relationships*; and she is at the beginning of the algebraic thinking development.

In conclusion, we agree with Davydov and Blanton, who suggest that non-numerical contexts can and should be employed with 4-5-year-old children to initiate the development of algebraic thought. We hope that the theoretical frame we proposed will help investigate the developmental routes of algebraic thinking in young students.

References

Blanton, M., Otálora, Y., Brizuela, B. M., Gardiner, A. M., Sawrey, K. B., Gibbins, A., & Kim, Y. (2018a). Exploring Kindergarten Students' Early Understandings of the Equal Sign. *Mathematical Thinking and Learning*, *20*(3), 167-201.

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

Boily, M., Polotskaia, E., Lessard, G., & Anwandter Cuellar, N. S. (2020). Les suites non numériques et le potentiel de la pensée algébrique chez les élèves du préscolaire. *Nouveaux cahiers de la recherche en éducation*, 22(1), 11-35.

Davydov, V. V. (1982). Psychological characteristics of the formation of mathematical operations in children. In T. P. Carpenter, J. M. Moser, & T. A. Romberg (Eds.), *Addition and subtraction: Cognitive perspective* (pp. 225–238). Lawrence Erlbaum Associates.

Davydov, V. V. (2008). Problems of developmental instruction. A theoretical and experimental psychological study. Novapublisher.

Eriksson, H., & Eriksson, I. (2020). Learning actions indicating algebraic thinking in multilingual classrooms. *Educational Studies in Mathematics*, 106(3), 363-378.

Kaput, J. J. (2008). What is algebra? What is algebraic reasoning? In *Algebra in the early grades* (pp. 5-17). Lawrence Erlbaum Associates.

Kieran, C. (2018). *Teaching and Learning Algebraic Thinking with 5- to 12-Year-Olds*. ICME-13 Monographs

Mellone, M., Ramploud, A., Di Paola, B., & Martignone, F. (2018). Cultural transposition: Italian didactic experiences inspired by Chinese and Russian perspectives on whole number arithmetic. *Zdm*, 51(1), 199-212.

Papadopoulos, I. & Patsiala, N. (2019). When the "Tug-of-War" game facilitates the development of algebraic thinking. *International Journal of Science and Mathematics Education*, 17(7), 1401-1421.

Polotskaia, E., Anwandter-Cuellar, N., & Savard, A. (2019). *La réussite en mathématiques au secondaire commence à la maternelle: Synthèse des connaissances sur les pratiques d'enseignement des mathématiques efficaces à la maternelle et au primaire pour réussir l'algèbre du secondaire.* Montréal, QC.

Radford, L. (2011). Grade 2 Students' Non-Symbolic Algebraic Thinking. In J. Cai & E. Knuth (Eds.), *Early Algebraization* (pp. 303-322). Springer.

Tremblay, S., Polotskaia, E., & Passaro, V. (2021). Réflexion autour du rôle du symbolisme littéral dans le développement de la pensée algébrique au primaire. *Revue québécoise de didactique des mathématiques*, *2*, 78-109.

INQUIRY MATHEMATICS TEACHING IN A UNIVERSITY BRIDGING COURSE: CHALLENGES FOR STUDENTS AND TEACHERS

Despina Potari*, Nikolaos Metaxas*, Barbara Jaworski**, and Theodossios Zachariades*

*Kapodistrian University of Athens, Greece; **Loughborough University, UK

This paper reports a study on the design and enactment of inquiry teaching approaches in a University Bridging Course, offered in a lecture format. In particular, the development of these approaches by the teachers, and the students' reactions to these as well as the tensions that both teachers and students experienced are investigated. The data consists of recordings of the lectures and students' work as well as interviews with the students and reflective discussions between the teachers. The analysis is based on the three layers of inquiry model of Jaworski (2019) and the framework of Potari et al., (2023). The identified tensions of teachers and students reveal challenges in the developmental process of inquiry approaches in the socio-cultural context.

INTRODUCTION

The last two decades university mathematics and its development has been the focus of research both in lectures and in group work settings (Dreyfus, Tabach & Rasmussen, 2023; Virman, 2021). In particular, ways of engaging students more actively in doing mathematics has been explored and inquiry-based learning approaches have been adopted even in quite advanced courses (Johnson et al., 2013). These approaches have been implemented in small group settings, workshops or tutorials, but rarely in lectures that are often characterized as "chalk and talk" where the lecturer presents the content following a specific structure with very little interaction with the students. However, there are studies that claim that even in the lecture format, lecturers may address students' needs in different ways (Petropoulou et al., 2020). Potari et al., (2023) illustrated tensions and contradictions that a lecturer faces in addressing demanding mathematics content to all students under the institutional and social conditions in which teaching was taking place. Mesa et al., (2020) point out that inquiry teaching, even in small groups, is difficult for the university teachers who also face several tensions related to their obligations to institutional factors as well as to mathematics. Moreover most studies focus either on the teachers or on the students while few attempts have been made to study university mathematics teaching both from the students' and teachers' perspectives (see PLATINUM project in Gómez-Chacón, et al. 2021). In this paper, we report a case study, of two university teachers who are also mathematics education researchers and use inquiry approaches in lectures. We use the three layers of inquiry process of Jaworski (2019) to address two research questions: (RQ1) How can inquiry approaches in the setting of a lecture be initiated? and (RQ2), What are the challenges for teachers and students in the process of engaging in inquiry?

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. 4, pp. 17–24). PME.

THEORETICAL BACKGROUND

Several definitions and terms have been used about inquiry teaching. Laursen and Rasmussen (2019) use the term Inquiry-Based Mathematics Education (IBME) as an umbrella to include several orientations related to inquiry approaches in teaching and learning mathematics at the university level. They consider this as a common vision that is based on four pillars, two emphasizing student behaviours, "engage deeply with coherent and meaningful mathematical tasks; "collaboratively process mathematical ideas" and two emphasizing instructor bevaviours "inquire into student thinking; foster equity in their design and facilitation choices" (p. 139). Jaworski (2019) embeds inquiry-based practice into the development of mathematics teaching through the three-layer model of inquiry. In the first layer the focus is the engagement of students in mathematical inquiry with the support of the teacher/lecturer. In the second layer the teachers engage in inquiry to scrutinize the students' activity and learn about creating learning opportunities for them. In the third layer, the teachers in collaboration with researchers are engaged in developmental research where they theorize on inquiry teaching and its development. In the current study, the layers of inquiry of Jaworski are considered more appropriate than the pilars of Laursen and Rasmussen to address both students' and teachers/researchers' perspectives and gain insights into teaching and learning. However, relating both frameworks could be a promising line of research.

Studies have indicated that inquiry approaches and their impact on students have been framed by the institutional and social conditions in which teaching takes place (Mesa et al. 2023). By considering the sociocultural context, Potari et al. (2023) use the Teaching Triad (TT) construct embedded in Activity Theory to identify certain relations between the elements of the TT, Mathematical Challenge – MC (mathematical inquiry, concepts, mathematical practices), Sensitivity to Students -SS (students' cognitive, affective and social needs) and Management of Learning – ML, (tools, learning environment, teaching actions) and to interprete these relations reflecting on the activity in which a lecturer participates. In that study, through the analysis of observations of teaching and using the TT, it appears that the lecturer's activity embodies affective and social sensitivity but rather little cognitive sensitivity and mathematical challenge. Interpreting these relationships on the basis of the lecturer's thinking in the discussions/interviews emerging contradictions were addressed and their role in the development of teaching.

In this paper, the two university teachers teach together a bridging course for the first year students. Both are mathematics education researchers trying to engage students in mathematical inquiry. They seek to change existing norms where teaching is in the amphitheatre, transmitted by the teacher with very little interaction between teacher and students. They also study students' activity to get feedback on the impact of their teaching approaches on their students. So, students' perspectives are also brought into consideration in the study of the inquiry process and the development of teaching. To bring some changes towards inquiry teaching approaches creates tensions (challenges) for the teachers and for the students and possibly for the learning outcomes of the students. As in the study of Potari et al. (2023), these tensions will be characterized in relation to specific interactions between teaching and learning using the TT and will be addressed in reflecting on the activity of the lecturers and of the students.

METHODOLOGY

Context

Our study took place within a series of lessons of an obligatory mathematics bridging course during the spring semester of the 2022-23 academic year at a mathematics department in Greece. The department offers this course for first year students in both winter and spring semesters and includes two 2-hour lessons every week for 13 weeks. The particular spring course was addressed to students who had not passed the winter course or had not enrolled in it during the winter semester. The spring semester course had an enrollment of about 100 students, but the attendance was only about 30 students. The course during both semesters had to cover the same content, which was basic set theory, binary relations, natural and integer numbers and cardinality of countable and uncountable sets. Most mathematics courses offered in the department are lecture-based, where the lecturer writes on the board and describes her reasoning.

Inquiry teaching approaches

The two university teachers (a male and a female) were mathematics education researchers who had not previously taught the course. They adopted inquiry approaches that have been discussed in our theoretical background. After the course began, feedback from students and conversations between them during reflective debriefings with each other helped to fine-tune their teaching decisions. They used the standard textbook which has been offered to students for the last five years and every week students were assigned as homework some tasks from the textbook, modified or completely new ones. Each homework contained between two and four problems and every week the teachers selected and discussed anonymously some of the students' submitted solutions in class. There were also two hourly problem sessions (third and 11th week) where students worked in groups to solve problems and then presented their solutions to their peers. During each group presentation~~s~~, the teachers encouraged students to ask questions, raise objections or request clarifications. There were also two 1-hour flipped classrooms (during fifth and tenth week) during which students in groups had to prepare and teach a specific theorem from the textbook. In the lectures, the teachers also emphasized the use of examples and diagrams; conjecturing and discussing the key ideas of proofs; eliciting students' ideas through questioning; encouraging students to share their ideas in the lecture; using power point presentations to discuss the main points of the lesson in parallel with the use of board.

Data and Data Analysis

Data included video recordings of lectures and group problem solving, audio recordings of meetings between the teachers and copies of student work generated over the semester. They also include students' responses on two online questionnaires regarding their views on the teaching methods. Finally, online interviews with four volunteered students were conducted lasting approximately one-hour each that were audio-recorded and transcribed. The work is in process so initial analysis of parts of the data has been reported in this paper adopting grounded theory techniques in episodes related to the main inquiry approaches that were enacted.

RESULTS

We structure our results below in relation to the two research questions.

Inquiry teaching approaches from teachers' and students' perspectives (RQ1)

The teachers implemented an inquiry-based learning approach to teaching where MC was promoted mainly through open tasks and questions as well as with an emphasis on mathematics concepts and proving processes. MC was balanced with SS in an attempt to make content accessible through the use of diagrams and examples and the feedback provided on homework (cognitive sensitivity); the encouragement of students to express their ideas and queries (affective sensitivity); and consider students who could not attend the lectures through the provision of a variety of resources (social sensitivity).

Studying the process of designing and enacting inquiry teaching from the teachers' perspectives, we can see that the teachers used resources from research in mathematics education and from their inquiry into students' and teachers' activity. For example, the transition from the distributive law of numbers to the sets in the group work activity came from reading the paper of Gabel & Dreyfus (2022) and discussing it through email: "It would be good to ask them to work in groups on a task similarly to the paper, to ask them first about the distributive law with numbers. Then to formulate this law in sets and to prove it explaining why in the sets both hold (Teacher B – 13-02-23)". Inquiring also into students' work provided teachers ideas about linking MC and SS. For example, Teacher A through the analysis of students' solutions in the homework tasks identified students' tendency to use definitions in proving: "I am interested in why students prefer the use of definitions and not of the theorems in proving. It is important to discuss this with them in relation to the tasks they tackle" (Teacher A – 22-02-2023). In the after class discussions the teachers also considered ML issues in seeking to increase students' participation and show SS: "I will keep some notes to think about "Why they do not participate". When you ask them some questions, you may need to give them time to think. Everybody could write something and then we ask someone to report." [Communications are translated from the Greek.]

How students experienced these inquiry teaching approaches was analysed through their responses in the questionnaires and the interviews. The students found especially helpful the resources provided (video-lectures, ppt presentations and the set of exercises) and they also valued the discussion in the amphitheater and the use of diagrams for supporting their understanding. The group work, the flipped classroom and the discussion on their solutions on the homework task were approved by all the students who were engaged in these. We provide below some examples of students' views about the inquiry teaching approaches adopted in the course:

" I liked a lot the flipped classroom. It helped us a lot to understand what we read, to understand it in depth, completely. If this could be done in all the course then we would understand the course completely 100%. There would be no confusion. But it takes time so it cannot be done in all the course". (flipped classroom)

"I was feeling more comfortable to participate comparing to the course in the previous semester, maybe because we were fewer students and I have been enculturated… In the previous semester, I was coming, observing and then leaving". (interaction in the lecture)

Challenges of the teachers and the students (RQ2)

Both the teachers and the students were confronted with difficulties stemming from different expectations and modes of working in school and in the other courses of the students and culminating in the creation of tensions. We exemplify some of these challenges in the group work activity and in the flipped classroom.

Group work activity: Below we give an excerpt from one group of four students working on a task to examine if there is a distributive property in sets similar to the distributive property of numbers. The role of union and intersection is left as an open question for the students to negotiate. The distributive law for numbers was chosen to engage the students in discussing about a property familiar to them from school while at the same time leading to an interesting investigation of its transformation to the sets.

1 Student A: Even for us that we started intuitively, there many ways to represent three sets and how they intersect each other … I don't think Venn diagrams are sufficient. I don't know, if intuition is a valid way to start.

2 Teacher B: So, your question is whether intuition is a safe way to start proving something?

3 Student A: I mean if it is worth it to start in this fashion, we could start formally if x belongs to the set on the left etc.

4 Teacher A: What do the others think about it?

5 Student B: Basically Venn diagrams are not proofs ,.., we need them for explaining, as an aid to get to the formal proof.

6 Teacher B: Why it's not a proof?

7 Student B: It's not a proof because we use the diagrams intuitively, we don't employ the definition to be driven to a valid claim. It's also written in the notes of the course.

8 Teacher B: If we check every case with a Venn diagram and a certain claim holds, is it a proof that the claim holds in general?

9 Student A: The answer is no, but I don't know.

10 Student B: In Venn diagrams we draw our own particular case on paper, on the other hand there could be other cases … so Venn diagram doesn't include every case while taking the definition we cover every single case.

11 Teacher A: Let's start the proof, as first step you can write x belongs to this set $A \cap (B \cup C)$, if and only if since we employ the definition, it maintains the equivalence relations.

In the excerpt above, we see that Teacher B tried to introduce students into an inquiry dialogue by making inquiry and deliberation dialogical moves (lines 4, 6 and 8) that are characterised by a high degree of MC . SS is indicated by Teacher B's encouragement of the students to exchange ideas (line 4) and by building on their epistemological views on proof production (lines 5 and 7). In line 11, Teacher A shifts the discussion in the construction of the formal proof reducing the MC. Here the tension for the teachers and especially for Teacher A as it was pointed out in the after class discussion concerned whether they should have had let the discussion go on and perhaps postpone the proof for the next session. Although both teachers concluded that in cases like this they had to sacrifice discovery for the purposes of the course, the question on how to achieve both goals had remained open. Moreover, the relationship between intuition and proof implies a contradiction (see Stouraitis et al. 2017) that from the teachers' point of view is not easy to handle while the students also feel uncertain about their role in proving (lines 7 and 9). Students also in the interviews considered on one hand their engagement in this form of inquiry in group work activities very helpful but on the other hand time consuming avoiding the solution of more exercises:

"Certainly it helps because you are in the position to solve the exercise, you discuss with the others, you listen to other ideas as well, other ways of thinking that you may have not thought them on your own, but because we are first year students (meaning they are not used to this approach), it may take a long time so we may not do many proofs or exercises in the lesson."

Flipped classroom activity: In the first flipped classroom activity the students presented a proof of the following proposition: *Let f: A→B function. Then f is onto if and only if it has a right inverse*. The following excerpt comes from the discussion after students' proof presentation.

1 Teacher B: Would you add anything to the proof of the textbook in order to make it more explanatory?

2 Student C: Well, in the proof there was a point where we have to notice that for every x in A, y=f(x) belongs to f(A), the main point of the proof and is a kind of a logical jump. The textbook says we notice that, but it is irrelevant.

3 Teacher B: How then could someone write the proof to make it more explanatory?

4 Student D: I think with the help of the Venn diagram, we have the two cases, either y is in f(A) or y is not. So, we notice that for every x in A then y which is f(x) is in f(A). A diagram would be more convincing here.

Students' narrative was an exact presentation of the textbook's proof without developing a kind of agency attached to it. On the contrary, the teachers' expectation was the exposure of the main strategic moves of the proof and an explanation of them: "Maybe we could have asked them to make a proof themselves. The process of proving was missing in the thinking about the proof. The focus was given to the structure although they made connections" (Teacher A).This resulted to two incommensurable discourses, which culminated to the tension between teachers' aims and students' responses that reflected on a textbook approach of the proof. A tension for the teachers was also to give on one hand the lead to the students to present their work and on the other to take care of the understanding of the other students in the class.

Overall, all the students in the interview recognised the importance of inquiry at different activities of the course (group work, flipped classroom) and the goals and actions of the teachers seem to meet students' needs. There is a tension of time for students and teachers as well (how much time to work in groups rather than being instructed). Students were interested in "reading" solved exercises uploaded in the course platform and also to have more exercises to solve on the board by the teachers: "I think that it is good to solve as many exercises as possible in the class, but this is not realistic. Maybe it would be better to solve the most important in the class and have many solved exercises in the e-platform". To solve more exercises in the class was often in contrast to the inquiry approaches that teachers aimed to engage the students. To be successful in the course final examination was a priority for the students that this required the practice of many exercises similar to those they had to face in the exams.

CONCLUDING REMARKS

Initiating an inquiry teaching approach at the university lectures is not an easy task. To interact with the students during the lesson may mean that only the "good" students participate in the discussion (see Petropoulou et al. 2020). In this study, we tried to engage students in different forms of inquiry such as discussing students' homework in the lecture, arranging problem solving group work, providing resources and flipped classroom activities to meet students' diffent cognitive, affective and social needs. The three layers of inquiry of Jaworski (2019) offered the opportunity to the university teachers who were also researchers in mathematics education to develop and enact inquiry activities for the students by co-designing and co-teaching (first layer) and through inquiring students' activity as well as their own activity as teachers and researchers (second and third layers) to understand how their decisions and students' interactions could facilitate learning in the lecture through experimentation and reflection. Engaging students in mathematical inquiry (MC) relevant to their needs (SS) created tensions for the teachers mainly stemming from their attempts to balance institutional constraints and students' expectations with their goals often initiated from their activity as researchers. Students seemed to align with the inquiry approaches but they were also concerned with success in the exams, expected to require less inquiry and more proving and solution of exercises. Through the ongoing and the retrospective analysis of students' homework, questionnaires and the interviews the university teachers became more aware of these tensions and developed ways to handle them.

References

Dreyfus, T., Tabach, M. & Rasmussen, C. Bafflement in an Inquiry-based College Mathematics Classroom. *Int. J. Res. Undergrad. Math. Ed. 9*, 557–569 (2023).

Gabel, M. & Dreyfus, T. (2022). Creating a Shared Basis of Agreement by Using a Cognitive Conflict. In: Biehler, R., Liebendörfer, M., Gueudet, G., Rasmussen, C., Winsløw, C. (eds)

Practice-Oriented Research in Tertiary Mathematics Education. Advances in Mathematics Education. Springer, Cham.

Gómez-Chacón, I. M., Hochmuth, R., Jaworski, B., Rebenda, J., Ruge, J., & Thomas, S. (2021). Inquiry in university mathematics teaching and learning: The PLATINUM project. Czech Republic: Masaryk University Press.

Jaworski, B. (2019). "Inquiry-Based Practice in University Mathematics Teaching Development". In *International Handbook of Mathematics Teacher Education: Volume 1*. Leiden, The Netherlands: Brill.

Johnson, E. (2013). Teachers' mathematical activity in inquiry-oriented instruction. The *Journal of Mathematical ehavior, 32*(4), 761–775.

Laursen, S.L & Rasmussen, C. (2019). I on the Prize: Inquiry Approaches in Undergraduate Mathematics. *Int. J. Res. Undergrad. Math. Ed. 5,* 129–146.

Mesa, V., Shultz, M. & Jackson, A. (2020). Moving away from lecture in undergraduate mathematics: Managing tensions within a coordinated inquiry-based linear algebra course. *International Journal of Research in Undergraduate Mathematics Education, 6,* 245-278.

Petropoulou, G., Jaworski, B., Potari, D. & Zachariades, T. (2020). Undergraduate mathematics teaching in first year lectures: Can it be responsive to student learning needs? *International Journal of Research in Undergraduate Mathematics Education, 6,* 347–374.

Potari, D., Jaworski, B. & Petropoulou, G. (2023). Theorizing university mathematics teaching: the Teaching Triad within an Activity Theory perspective, *Educational Studies in Mathematics, 114*:1–16, https://doi.org/10.1007/s10649-023-10244-x

Stouraitis, K., Potari, D. & Skott, J. (2017). Contradictions, dialectical oppositions and shifts in teaching mathematics. *Educational Studies in Mathematics, 95*(2), 203–217.

Viirman, O. (2021). University mathematics lecturing as modelling mathematical discourse. International *Journal of Research in Undergraduate Mathematics Education, 7*(3), 466-489.

PSYCHOLOGICAL THEORY AND INNOVATIONS IN REFORMS OF MATHEMATICS EDUCATION – A QUESTION OF DISCOURSE AND GRAMMAR

Johan Prytz[1], Uffe Thomas Jankvist[2], Linda Ahl[1], and Iresha Ratnayake[1]

[1]Uppsala University, [2]Aarhus University

This philosophical essay delves into the role of theories in mathematics curriculum reforms, particularly how theories can contribute to creating and implementing innovations. Using the concepts of discourse and grammar of schooling, we investigate two well-researched Swedish curriculum reforms. With these two concepts, we discuss the possible contribution of the underlying theories to the success of one reform and the failure of the other.

INTRODUCTION

This philosophical essay concerns the role of psychological theory (PT) in mathematics curriculum reforms and how that role can be conceptualized and analyzed. We aim to discern how PTs can provide innovations and influence formal curriculum documents, teaching materials, and teachers in different, more or less efficient, ways. To study different ways of influence, we use the concepts of *discourse* and *grammar of schooling*. The former is well-known in mathematics education research, while the latter is not. In brief, each concept concerns a structure that can exist in an educational text simultaneously, which can affect teachers. To illustrate what is more or less efficient, we study two cases. One case was successful in the sense that key innovations in the formal curriculum had an impact on textbooks, teachers, and student results. The other case was less successful because key innovations in the formal curriculum did not impact teachers and textbooks, and student results did not improve. The two cases are the Swedish mathematics curriculum reform of 1980 (success) and the subsequent Swedish mathematics curriculum of 1994 (not a success). Our guiding questions (GQ) have been the following:

GQ1: How were the key innovations of the two cases connected to PT?

GQ2: In what respect did the innovations and PT contribute to discourses and grammar of schooling?

We draw on a synthesis of results from previous research concerning the two reforms mentioned above. This paper's original contribution is a comparison of the two reforms and an application of the concepts of *discourse* and *grammar of schooling.*

PREVIOUS RESEARCH

The design of a formal curriculum, that is, documents issued by a governmental body intended to steer teachers and their teaching, can positively affect what teachers teach

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. 4, pp. 25–32). PME.

and what students learn (Schmidt et al. 2001; Schmidt & Prawat, 2006; Prytz, 2020). In brief, it is a matter of content covered and emphasized in the formal curriculum that receives more attention in teaching and that students learn more about that content. However, covering or emphasizing content can be done in different and more or less efficient ways (Schmidt & Prawat, 2006; Prytz 2020).

There are thus different ways to introduce an innovation based on a PT in mathematics education through the formal mathematics curriculum (i.e., having it in the text). We seek to understand these ways through the *discourse* and *grammar of schooling*. Innovation is a general concept here and includes something new compared to the previous curriculum. It can be new concepts or the design of guidelines.

Discourse is a well-used concept in current research about curriculum reforms in mathematics education. In a recent comprehensive research overview on mathematics curriculum reforms (Shimizu & Vithal, 2023a), discourse sometimes refers to debates or discussions going on in society that influence the formal curriculum (e.g., Shimizu & Vithal, 2023b; Gosztonyi et al., 2023). Sometimes, discourse refers to more specific language structures and psychological or psychosocial aspects of the people participating in the discourse. For instance, Ruiz et al. (2023) consider mathematical discourse to influence how people understand and reason about mathematics. Quirke et al. (2023) view discourse as influencing teachers’ professional identity. Pinto and Cooper (2023) see discourse as an essential part of mathematics teachers’ identity (i.e., how mathematics is taught and what the best source of knowledge about teaching is). Both Quirke et al. (2023) and Pinto and Cooper (2023) point out that other groups than teachers (stakeholders with an interest in mathematics education) can belong to other discourses. Discursive differences can pose a problem if, for instance, the discourse of politicians and school bureaucrats formally deciding the curriculum content is different from the teacher discourse. More precisely, it can hinder the formulation and implementation of a formal curriculum due to varying understandings of expressions and concepts and even conflicting expectations.

Hence, the research into the psychological or psychosocial aspects of discourses about mathematics education provides important insights into the design of a formal curriculum and the effects it can have on teachers.

The ‘grammar of schooling’ concept concerns teachers’ need for routines for managing students in time and space. The term is coined by the historians of education Tyack and Cuban (1995) in their seminal book *Tinkering toward Utopia: century of public school reform*. In this book, and with the US as the case, they seek to explain why educational reforms in the 20th century have succeeded or, most often, not succeeded. The story begins already in the 19th century, when Tyack and Cuban (1995) observe that several reforms, when mass education expanded quickly, were highly successful, for instance, the introduction of age-graded classes, schedules built up by mandatory subjects, and teaching practices adapted to this type of organization. These innovations spread rapidly and became basic structures of daily life in almost all schools – they became a grammar of schooling. This grammar became stable. Several reforms of the 20th century, not seldom labeled 'progressive', have aimed to change these structures fundamentally, but they all failed. One explanation of the latter is that teachers liked the established grammar. It:

> […] enabled teachers to discharge their duties in a predictable fashion and to cope with the everyday tasks that school boards, principals, and parents expected them to perform: controlling student behavior, instructing heterogeneous pupils, and sorting people for future roles in school and later life. (Tyack & Cuban, 1995, p. 86).

Other explanations of failed reforms are 1) loss of legitimacy as the reformers lost contact with the general public's conception of schooling and 2) burnout and overload among reformers or teachers since creating new routines became overwhelming (Tyack & Cuban, 1995).

The concept of grammar of schooling is uncommon in mathematics education contexts. The term is not used in the overview of mathematics curriculum reforms (Shimizu & Vithal, 2023a). The concept does occur implicitly in some chapters when the need to support teachers with materials and pay attention to teachers' situations is stressed. Nevertheless, the teacher's need to manage students in time and space is not highlighted as a prominent factor or a quality of support.

When we seek to understand a reform process with the concepts of grammar of schooling and discourse, we partly consider the same things in curriculum documents and textbooks, for instance, what the mathematical content was and how it was described. However, different aspects are brought to the fore. Grammar of schooling highlights teachers' need for routines for managing students in time and space, whereas discourse concerns teachers' language, identity, and social relations. In this paper, we illuminate how analyses of mathematics curriculum reforms based on discourse and grammar of schooling differ and how they can be combined. We also illuminate the benefits of using grammar of schooling rather than just discourse.

THEORY AND METHOD

In our analysis, a psychological theory is a set of concepts and assumptions about some aspect of the physical or the ideational worlds. In a scientific theory, questions are central, involving an explicit methodology about how to pose and answer questions. Our definition of a psychological theory (PT) is based on the Oxford English Dictionary definition of psychology as "the scientific study of the nature, functioning, and development of the human mind, including the faculties of reason, emotion, perception, communication, etc." (OED 2024).

Our study involves two PTs: George Miller's theory on working memory (Miller, 1968) and the NCTM mathematical competence framework, also known as process standards (NCTM, 1989). Their status as scientific are different, but both are psychological. Miller's theory focuses on how the human mind functions, particularly the processing of information and limits of the working memory. The NCTM framework concerns the nature of the human mind when it comes to mathematical thinking and how it consists of different competencies.

A PT can exist in a formal curriculum document. When we talk about PT existing in a curriculum document, it is a matter of expressions or concepts central to the PT appearing in the text.

In our analysis of how curriculum documents can influence teachers and their teaching, we have considered two types of influence: *discourse* and *grammar of schooling*. By *discourse,* we mean how people in a particular context speak about a topic and how this way of speaking influences peoples' self-perception, their perception of other people, and their perception of the world around them. By *grammar of schooling,* we mean school structures that concern how teachers and students are positioned and what they do in time and space.

As mentioned, this essay is a synthesis study, bringing together results from previous studies, mainly our own, in a new way. These results come from analyses of reports concerning the two curriculum reform programs, previously published analyses of impact on textbooks series, and analyses of national and international assessment results.

ANALYSIS

The period from 1980 to 1995 was successful for mathematics education in Sweden; it is the best record so far, at least if student results in national and international tests are considered, as these improved significantly. Interestingly, the same formal curriculum was in effect for much of this period: the 1980 national curriculum, which was replaced in 1994 (Prytz, 2020).

The 1980 mathematics curriculum marked a definitive break with the New Math movement and consolidated the Back to Basics movement, a shift going on since the mid-70s. The Back to Basics movement meant that arithmetic was emphasized at the expense of algebra and geometry, and computational procedures were favored over understanding concepts. Problem-solving was emphasized, but then in the context of everyday applications. Moreover, diagnostic materials were further underscored as an important tool for teachers (Prytz, 2018, 2020).

The 1980 mathematics curriculum also marked a shift in connection to science. The connection to the scientific discipline of mathematics became weaker as concepts from set theory were not used to organize school mathematics. Instead, psychology became the scientific domain to inform teachers how to teach.

In the 1970s, the Swedish national school authority supported two projects to improve elementary mathematics skills, particularly arithmetic. The background was low student results, especially in arithmetic, and the New Math reform, launched in 1969, was identified as a possible cause (Prytz, 2020).

The most extensive project of the two was PUMP (*Processanalyser av undervisning i Matematik/Psykolingvistik*, [Process analysis of teaching in mathematics/psycho-linguistics]). PUMP was a four-year research and development project focused on arithmetic and school years 1 to 6. The scientific theory guiding the work was George Miller's theory on working memory and cognitive load, i.e., a theory from psychology (Prytz et al., 2022).

In PUMP, it was discovered, through classroom observations and textbook analysis, that the teaching progression in arithmetic was too fast. The students encountered too complicated exercises too early. In terms of Miller's theory, the exercises involved too many new concepts and procedures, leading to cognitive overload and hampered learning. The PUMP people, therefore, sought to develop a detailed scheme of putting exercises in a sequence that did not lead to cognitive overload. They developed diagnostic material to determine what exercises the students should work on (Prytz et al., 2022).

Much of the PUMP thinking went into the 1980 mathematics curriculum and the commentary material: on the one hand, as specific and concrete guidelines about the type of exercises the students should encounter and their sequencing, and on the other hand, as references to reports from the PUMP project (Prytz, 2018, 2020). In comparison to the previous 1969 curriculum, this was an apparent innovation. We can also observe that popular textbook authors quickly picked up the new guidelines, even some years before the 1980 reform (Prytz et al., 2022).

There is further evidence that development work in the 1970s and the 1980 curriculum section about arithmetic positively affected teaching and learning. From SIMS1980 to TIMSS1995, the Swedish results improved considerably, particularly in arithmetic, in contrast to the areas of algebra, geometry, and statistics, where the improvements were more modest. Moreover, there had been no research and development projects in those areas, and the curriculum documents were briefer and more general as to the type of exercises the students should encounter and their sequencing (Prytz, 2020).

All these aspects of the 1980 curriculum reform can be seen as something that influenced a discourse about mathematics teaching and, eventually, teachers' professional identity and even their way of teaching.

From a grammar of schooling perspective, we can see that the PUMP project and the 1980 curriculum also changed the grammar of schooling. They provided new daily routines for teachers' management of students; more precisely, routines for allocating, in time, the right type of exercises for students to work with.

In 1994, the 1980 curriculum was replaced. The design of the new 1994 curriculum and the commentary material were, in certain respects, very different, even though the mathematical content areas (arithmetic, algebra, geometry, and statistics) were the same. The innovative feature was the emphasis on competencies. The newest in this respect was the reasoning, communication, and problem-solving competencies; they did not have a prominent and clear position as specific competencies in the 1980 curriculum (Boesen et al., 2014; Prytz, 2015). The scientific background of the emphasis on competencies was competence theories, i.e., models where the phenomenon of knowing is divided into different competencies. The influence came from the NCTM common core standards, which were being developed in the US. In the early 1990s, Swedish curriculum developers had good contacts with NCTM people (Helenius & Ahl, 2023).

Implementing the 1994 curriculum, particularly the innovative parts on reasoning, communication, and problem-solving competencies, did not work well. When Boesen et al. (2014) studied the degree to which teachers had implemented the 1994 curriculum, they found that teachers had difficulties understanding and teaching the parts about reasoning and problem-solving competencies, even though they were rather positive towards them. Instead, much of the teaching was focused on procedures. Another circumstance is that results in large-scale international assessments began to decline by the end of the 1990s.

This change in the formal curriculum 1994 can be seen as a change in discourse and an attempt to change teacher identity. Nonetheless, there is little to see when we apply a grammar of schooling perspective on the 1994 curriculum and the commentary material, related to the fact that the curriculum and the commentary material were significantly shorter and contained much fewer details and examples concerning what to teach and in which order. So, even though the curriculum emphasized the reasoning, communication, and problem-solving competencies, there was no support to create new routines – a new grammar – to teach the reasoning, communication, and problem-solving competencies. Our preliminary analysis of textbooks published around 1995 and onwards yields a similar result: the textbooks did very little to support teaching about communication, reasoning, and problem-solving (Prytz et al., 2024).

So, when our analyses of the 1980 and 1994 curriculum reform are compared, we have an example where the grammar of schooling can contribute to an explanation of why curriculum reforms, particularly the implementation of the innovative features, fail or succeed, which discourse does not, in this case.

CONCLUSIONS

In this paper, we have illustrated how PT can exist in curriculum documents in two ways: as parts of discourse or as parts of the grammar of schooling. We have also illustrated how PT can be part of both discourse and grammar in the same curriculum document (1980) and how PT can be part of just discourse (1994).

Our findings suggest that PT being part of both discourse and grammar, or just grammar, in the curriculum documents is vital for the successful implementation of innovations. Therefore, to be successful as a reformer in mathematics education, it is not sufficient just to win or dominate the discourse; you also need to create a grammar of schooling that matches that discourse. More precisely, creating teachable sequences of explanations, exercises, and other activities enables teachers to manage students in space and time.

Our findings also give reasons to think about PTs having different potentials to change the discourse and grammar of schooling in connection to curriculum reforms, especially regarding the creation and implementation of innovations. For instance, a central component in Miller's theory is that there is an upper limit to how much information the working memory can handle. That is a highly relevant condition to create teachable sequences of mathematical explanations and exercises, i.e., routines for teachers. In contrast, the NCTM competence framework does not have a corresponding component. So, in that respect, Miller's theory should have better potential of changing the grammar of schooling. On the other hand, Miller's theory concerns cognition in general. It will have little potential in mathematics education unless combined with some mathematical content. One can argue this happened in algebra concerning the Swedish 1980 curriculum reform. Teaching and results improved more in the area (arithmetic) where Miller's theory had been involved in the development work and not so much in algebra; the latter was "untouched" by Miller's theory. Similarly, Miller's theory, per se, due to its general character, should have little potential to influence discourse on mathematics education in contrast to the NCTM competence framework, which is specific to mathematics and covers the whole mathematics curriculum. Hence, both theories appear to have strengths and weaknesses concerning curriculum reforms. Our advice to reformers is to not lean on just one theory but several and then learn about their strengths and weaknesses. The latter is, of course, something researchers can dig into. In this paper, we have offered some concepts to consider.

References

Boesen, J., Helenius, O., Bergqvist, E., Bergqvist, T., Lithner, J., Palm, T., & Palmberg, B. (2014). Developing mathematical competence: From the intended to the enacted curriculum. *The Journal of Mathematical Behavior*, *33*, 72-87.

Pinto, A., Cooper, J. (2023). Boundary Crossing in Curriculum Reform. In: Shimizu, Y., Vithal, R. (eds) *Mathematics Curriculum Reforms Around the World*. Springer.

Gosztonyi, K., van den Heuvel-Panhuizen, M., Makinae, N., Shimizu, S., van Zanten, M. (2023). International Co-operation and Influential Reforms. In: Shimizu, Y., Vithal, R. (eds) *Mathematics Curriculum Reforms Around the World.* Springer.

Helenius, O., & Ahl, L. M. (2023). *Hur bör man förändra kursplaner i matematik?* Svenskt Näringsliv.

Miller, G.A. (1968). The psychology of communication: seven essays. London: Penguin P.

NCTM (National Council of Teachers of Mathematics. Commission on Standards for School Mathematics) (1989). *Curriculum and evaluation standards for school mathematics*. Reston, Va.: The Council.

Prytz, J. (2015). Swedish mathematics curricula, 1850–2014. An overview. In Bjarnadóttir, K., Furinghetti, F., Prytz, J., & Schubring, G. (2015). *"Dig where you stand" 3:*

Proceedings of the Third International Conference on the History of Mathematics Education. Uppsala University.

Prytz, J. (2018). The New Math and School Governance: An Explanation of the Decline of the New Math in Sweden. In: Furinghetti, F., Karp, A. (eds) *Researching the History of Mathematics Education. ICME-13 Monographs*. Springer.

Prytz, J. (2020). Framing for success: Governance of Swedish school mathematics, 1980–1995. *Nordic Journal of Educational History*, *7*(1), 3–32.

Prytz, J., Ahl, L. M., & Jankvist, U. T. (2022). An Innovation's Path to Mathematics Textbooks: A Retrospective Analysis of the Successful Scaling of the Swedish PUMP Project. *Implementation and Replication Studies in Mathematics Education*, *2*(2), 241–288.

Prytz, J. Ratnayake, I. Ahl, L. M., & Jankvist, U. T. (2024). Go with the flow – a national professional development program for mathematics teachers in Sweden, 2013–2016. Paper to be presented at the Tenth Nordic Conference on Mathematics Education in June 2024 at Aarhus University, Campus Emdrup in Copenhagen, Denmark.

Schmidt, W.H., McKnight, C. C., Houang, R. T., Wang, H., Wiley, D. E., Cogan, L. S., & Wolfe, R. G. (2001). *Why schools matter: a cross-national comparison of curriculum and learning*. Jossey-Bass, A Wiley.

Schmidt, W.H., & Prawat, R. S. (2006). Curriculum coherence and national control of education: issue or non-issue?. *Journal of curriculum studies, 38*(6), 641–658.

Shimizu, Y., & Vithal, R. (2023a). Mathematics curriculum reforms around the world: The 24th ICMI study. Springer.

Shimizu, Y., Vithal, R. (2023b). School Mathematics Curriculum Reforms: Widespread Practice But Under-Researched in Mathematics Education. In: Shimizu, Y., Vithal, R. (eds) *Mathematics Curriculum Reforms Around the World.* Springer.

Tyack, D. & Cuban, L. (1995). Tinkering toward Utopia: century of public school reform. Harvard University Press.

Quirke, S., Espinoza, L., Sensevy, G. (2023). Teacher Professional Identity and Curriculum Reform. In: Shimizu, Y., Vithal, R. (eds) *Mathematics Curriculum Reforms Around the World*. Springer, Cham.

SELF-EFFICACY EXPECTATIONS OF MATHEMATICS UNIVERSITY STUDENTS

Stefanie Rach[1], Timo Kosiol[2], and Stefan Ufer[2]

[1]OVGU Magdeburg, [2]LMU Munich

Self-efficacy expectations, which are learners' estimation of being able to solve a task, are an important motivational variable in learning processes. Learners with high expectations may be more ambitious when dealing with mathematical tasks, particularly in the challenging entry stage of a university program. It is not clear how situation and person characteristics influence these expectations. Results of a study with 338 students enrolled in mathematics study programs show that stable person characteristics, such as different facets of self-concept, and the mathematical practice required in the task (calculating, modelling, and proving) interact in predicting self-efficacy expectations. The results shed light on the complex interplay of person and situation characteristics, highlighting the situation-specificity of expectations.

INTRODUCTION

At the transition from school to university mathematics, challenging mathematical learning tasks are offered for students' self-regulated learning. However, many students tend to deal superficially with these tasks. It is assumed that persons with high self-efficacy expectations invest more time on tasks because they can overcome temporal complications and impasses. Thus, high self-efficacy expectations are of particular importance for successful learning processes in this context. We assume that person characteristics, such as mathematical self-concept, as well as situation characteristics, such as the mathematical practice foregrounded in a task, influence the emergence of self-efficacy expectations as a situational measure. It is unclear however, how these specific characteristics interact in influencing self-efficacy expectations.

THEORETICAL BACKGROUND

Context: study entry phase

The transition from school to university is a challenging phase for students because in many countries not only the institution changes but also the predominant mathematical practices targeted by instruction. In university mathematics study programs, *proving tasks* predominate whereas in school more focus is on *calculation tasks* or *modelling tasks* in which mathematical contents are used to solve real-world problems (Engelbrecht, 2010; Rach & Heinze, 2011). Previous research shows that some students can anticipate which mathematical practices, e. g., modelling or proving, are central for school or university mathematics (Rach et al., 2014) and that they can report specific motivational tendencies towards these practices (Rach et al., 2017).

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. 4, pp. 33–40). PME.

Students experience failure (di Martino & Gregorio, 2018), in particular when working on challenging learning tasks which are often demanded as compulsory exercise in undergraduate mathematics courses. As a result, superficial task engagement strategies, such as copying task solutions from peers (Liebendörfer & Göller, 2016) without processing them deeply (for example, by self-explaining the copied solutions) are frequently observed. This superficial learning behaviour goes along with reduced learning gain (Rach & Heinze, 2011). Reasons for superficial processing have been sought in students' characteristics (Berger & Karabenick, 2011) as well as task characteristics (Schukajlow et al., 2012), such as the mathematical practices foregrounded in a task.

Self-efficacy expectations

Bandura (1977) understands "efficacy expectation" as "the conviction that one can successfully execute the behavior required to produce the outcomes" (p. 193). A similar construct "expectation of success" as "individuals' beliefs about how well they will do on an upcoming task" is used by Eccles and Wigfield (2020, p. 3) in the situated expectancy-value model which explains students' choices, persistence, and performance in learning processes. Both approaches conceptualize such expectations as a situational construct that is affected by more stable person characteristics, as well as characteristics of the learning situation, entailing the concrete learning task.

Empirical research in mathematics classrooms only partly supports the assumption that self-efficacy expectations depend on characteristics of the learning task. Schukajlow et al. (2012) reported only small differences in ninth graders' self-efficacy expectations between modelling tasks and word problems, and they could not identify any differences to intra-mathematical problems. In the study of Krawitz and Schukajlow (2018), ninth and tenth graders reported different self-efficacy expectations concerning tasks which foreground different mathematical practices. Self-efficacy expectations in mathematics are often measured using different mathematical tasks from different mathematical topics that foreground different practices (Parker et al., 2014). Yet, self-efficacy expectations are not differentiated based on such task characteristics despite the results of Krawitz and Schukajlow (2018).

Self-efficacy expectations, as a situational construct, are assumed to also depend on more stable person characteristics. Similar to self-efficacy expectations, self-concept describes a learners' image concerning their knowledge and skills in a certain domain. Both constructs share many similarities such as the focus on perceived skills, domain-specificity, and multidimensionality. Whereas self-efficacy expectations focus on the potential for solving a specific, given task in the (near) future, self-concept is framed as a relatively stable person characteristic, which arises from retrospective experiences (Marsh et al., 2019). Theoretical works and empirical works with undergraduate students underpin a strong relation between self-efficacy expectations and self-concept (Bong & Skaalvik, 2003; Pajares & Miller, 1994). Thus, for estimating whether oneself can solve a particular task, it is plausible that students draw on their image of their own domain-specific skills. When studying effects of self-concept, it is advisable to consider measures of actual domain-specific performance as a covariate to disentangle effects of students' actual skills and their own image of these skills.

In line with Eccles & Wigfield (2020), low self-efficacy expectations are one possible explanation for superficial task processing when dealing with challenging learning tasks. Determinants of self-efficacy expectations are still a matter of research especially when the mathematical practices foregrounding a task vary.

Research questions of the current study

In this contribution, we focus the joint influence of person and situation characteristics on students' self-efficacy expectations regarding learning tasks in university mathematics courses. In particular, we analyse (Q1) to which extent students' self-efficacy expectations differ systematically between persons and between tasks, (Q2) whether the foregrounded practice in the tasks (calculating, modelling and proving) contributes to variance explanation, (Q3) whether students' individual self-efficacy reports vary systematically over different foregrounded practices, and (Q4) how students' self-efficacy expectations for each practice relate to their self-concepts concerning the same three practices and their actual mathematical performance.

METHODS

This study is part of the project SISMa ("Self-concept and Interest when Studying Mathematics"). The sample comprises 338 first-year students of five different study programs (*n* = 92 general mathematics, *n* = 89 financial mathematics, *n* = 102 teacher education for the high attaining school track, *n* = 35 teacher education for other school types, *n* = 14 computer science, *n* = 6 missing) in two "Analysis 1" courses who voluntarily participated in this study.

To measure *self-efficacy expectations*, we used a questionnaire consisting of twelve tasks (four tasks for every practice, see also Rach et al., 2017) in the field of Analysis. Students reported their self-efficacy expectation for each task by evaluating the statement "I think I can solve this task" on a four-point likert scale from "I don't agree" (0) to "I agree" (3).

- Example task for calculating: "Let f be $f(x) = \frac{\sqrt{1+x}\cdot e^x}{4+x^2} - 1$. Calculate the extrema of the function f."
- Example task for modelling: "By metal, you should produce a cylindrical can with a prescribed volume. For which radius is the material consumption minimal?"
- Example task for proving: "Let $f\colon \mathbb{R} \to \mathbb{R}$ be a differentiable function. Show that f is continuous."

As personal variables, we measured *self-concept concerning calculating*, *modelling*, and *proving* with an approved questionnaire (Rach et al., 2019; calculating; 5 items, α = .71; modelling: 5 items, α = .80; proving: 4 items, α = 78, item example:

"Understanding mathematical proofs is easy for me."). We measured students' *actual knowledge of advanced mathematics* with an approved multiple-choice test (8 items, $\alpha = .58$). All personal variables were z-standardised for analyses and show small to medium correlations ($r < .35$).

Our study implements a cross-classified multilevel design because self-efficacy expectations vary over two (random) factors: the person ($N = 338$) and the task ($N = 12$). We estimated linear mixed models with *lme4* in R, version 4.3.0, since they are a sensible way to analyse such complex data (Bates et al., 2015).

RESULTS

(Q1) To investigate to which extent students' self-efficacy expectations differ between persons and tasks, we estimated a linear mixed model containing only both random effects. Random effect variances (table 1) indicate that self-efficacy expectations vary significantly between persons (33.5% variance of self-efficacy expectations explained, model 1) and, also significant but to a smaller degree, between tasks (3.5% variance of self-efficacy expectations explained, model 2).

	Model 1	Model 2	Model 3	Model 4
Random effect: person	*Var* = 0.25	*Var* = 0.25	*Var* = 0.25	*Var* = 0.31
Random effect: task		*Var* = 0.03	*Var* = 0.02	*Var* = 0.02
Random slope effect: person*practice				*Var* = 0.16[a] *Var* = 0.17[b]
Residual	*Var* = 0.50	*Var* = 0.47	*Var* = 0.47	*Var* = 0.42
Fixed effect: practice			$F(2, 9.00)=1.67$, $p = .24$	$F(2, 9.64)=1.60$, $p = .25$
Difference between models	-	$\chi(1) = 164.4$, $p < .001$ (Model 1 and 2)	$\chi(2) = 3.69$, $p = .16$ (Model 2 and 3)	$\chi(5) = 87.9$, $p < .001$ (Model 2 and 4)

Table 1: Results of linear mixed models with dependent variable self-efficacy expectations; *Var* = Variance. N = 3,982 observations; [a]difference between proving and calculating; [b]difference between proving and modelling

(Q2) The practices foregrounded in the tasks (calculating, modelling, or proving) do not contribute to the variance explanation in self-efficacy expectations as a fixed effect (model 3). Indeed, mean self-efficacy expectations, aggregated over all persons and all tasks per practice, are similar for all three practices: calculating ($M = 2.25$; $SD = 0.83$;

N = 1,325), modelling (*M* = 2.07; *SD* = 0.88; *N* = 1,322), and proving (*M* = 2.07; *SD* = 0.87; *N* = 1,335).

(Q3) To analyse whether students' individual self-efficacy reports vary systematically between different foregrounded practices, we included a random slope for the effect of practice over persons (model 4). This means that separate person intercepts are estimated for each practice and each student (like in a three-dimensional model).

(Q4) To analyse which person characteristics relate to self-efficacy expectations, we integrated these characteristics as fixed effects in model 4. All self-concept facets significantly predict self-efficacy expectations (calculating: $F(1, 330.9) = 25.1$, $p < .001$; modelling: $F(1, 331.2) = 12.3$, $p < .001$; proving: $F(1, 330.0) = 29.8$, $p < .001$) but there is no clear evidence that besides these perceptions, students' actual mathematical knowledge relates to self-efficacy expectations ($F(1, 331.0) = 3.66$, $p = .06$). Interaction effects show that the influence of the different self-concept facets on self-efficacy expectations depend substantially on the practices of the tasks, except for proving (table 2, last column). Trend analyses indicate that self-concept concerning proving predicts self-efficacy for all foregrounded practices. Further, self-concept concerning modelling relates stronger to self-efficacy regarding modelling tasks ($b = 0.16$) than the other two practices ($b = 0.08$ or $b = 0.05$). Self-concept concerning calculating shows a stronger prediction on self-efficacy regarding calculation tasks ($b = 0.18$) than to modelling tasks ($b = 0.09$), while the difference is non-significant between calculation and proving tasks ($b = 0.12$). Summarizing, students' self-efficacy expectations in task over all practices is significantly predicted by their proof-related self-concept. Beyond this, self-efficacy for calculating and modelling tasks are significantly predicted by the related self-concept facets. Interestingly, calculation-related self-concept significantly predicts self-efficacy on proving tasks.

		Self-efficacy expectations			
		Calculating	Modelling	Proving	Self-concept × practice
Self-concept	Calculating	0.18 [0.12; 0.25]	0.09 [0.03; 0.16]	0.12 [0.05; 0.19]	$F(2, 331.0) = 4.10$*
	Modelling	0.08 [0.01; 0.14]	0.16 [0.10; 0.23]	0.05 [-0.02; 0.11]	$F(1, 332.0) = 5.92$**
	Proving	0.13 [0.07; 0.20]	0.15 [0.08; 0.21]	0.22 [0.15; 0.28]	$F(1, 329.7) = 2.95$, $p = .06$

Table 2: Results of contrast analyses for predicting self-efficacy expectations of tasks with different practices by different self-concept facets; trend coefficients and confidence intervals for confidence level .95; *** $p < .001$, ** $p < .01$, * $p < .05$

DISCUSSION

Our main goal was to investigate how situation and person characteristics interact in their relation to students' self-efficacy expectations. The results unveil that learners' self-efficacy expectations vary stronger between learners than between tasks and their foregrounded practices (Q1). As at the transition from school to university, the mathematical practices, which predominate the learning process, change (Engelbrecht, 2010), it was plausible to assume that students' self-efficacy expectations differ by the foregrounded practices as a situation characteristic (Q2). The results do not indicate that students report lower self-efficacy expectations for proving tasks than for calculating or modelling tasks (similar results as reported in Schukajlow et al., 2012). However, students differentiate their expectations systematically between different foregrounded practices (Q3). In general, self-efficacy expectations seem to be more strongly influenced by the person or the interaction of person and learning situation than by the concrete situation itself. This indicates that one-size-fits-all task designs may be of little promise for boosting participants' self-efficacy expectations.

Adapting to learners' characteristics requires to understand the interaction of personal and situational variables in the genesis of self-efficacy expectations. Further analyses show that students' self-concept, rather than their actual mathematical knowledge, predict students' self-efficacy expectations (Q4, see Bong & Skaalvik, 2003). This indicates that a positively-realistic self-concept may trigger beneficial self-efficacy expectations, and potentially sustained task engagement. Developing such self-concept requires opportunities to perceive oneself as competent as it is proposed in the self-determination theory of Ryan & Deci (2020). Differentiating the analyses by practices unveils that – as one would expect – self-concept regarding a specific practice is predictive for self-efficacy expectations on tasks foregrounding this practice. We also find that self-concept regarding proving predicts self-efficacy expectation comparably strongly for all practices foregrounded in the task. Possible explanations must be investigated in future research but one reason could be that students anticipate proofs to play a role for all learning tasks in the new university context. Finally, also self-concept regarding calculation significantly predicts self-efficacy for tasks foregrounding the other practices. For proving tasks, which are a major obstacle for many learners in the study entry phase, a positive self-concept regarding calculating, which is a practice that is familiar from school to many students (Rach et al., 2014), might benefit students' self-efficacy expectations and potentially engagement towards this challenging practice. This more practical conclusion offers levers to increase self-efficacy expectations by offering opportunities for positive, but still realistic skill perception also when working on other mathematical practices (e.g., calculating) than those posing specific problems (e.g., proving) in the study entry phase.

Our findings are limited by the fact that we only used tasks in the field of Analysis. Moreover, we developed the mathematical tasks for this study by differentiating between practices. Other tasks characteristics beyond these practices, such as the targeted mathematical concepts, e. g., limits or derivations, or explicit scaffolding approaches to support self-efficacy could be analysed in further studies. Finally, we could not investigate actual task engagement in this study. Even though, there is substantial evidence for the effects of self-efficacy expectations on task engagement (Berger & Karabenick, 2011), investigating the whole effect chain from person and situation characteristics over self-efficacy expectations to task engagement and potentially task performance or learning gain would be of substantial interest.

This study provides theoretical insights into the complex interaction of person and situation factors in the genesis of students' self-efficacy expectations in a university mathematics program. The results point to interesting interactions but more research is needed to understand how person and situation characteristics work together on the micro-level and how they influence students' learning-related decisions during task engagement. This call for analysing the situated nature of motivational constructs is in line with the analyses of Eccles and Wigfield (2020) and Schukajlow et al. (2023).

References

Bandura, A. (1977). Self-efficacy: Toward a Unifying Theory of Behavioral Change. *Psychological Review, 84*(2), 191–215.

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48.

Berger, J.-L., & Karabenick, S. A. (2011). Motivation and students' use of learning strategies: Evidence of unidirectional effects in mathematics classrooms. *Learning and Instruction, 21*, 416–428.

Bong, M., & Skaalvik, E.M. (2003). Academic self-concept and self-efficacy: how different are they really? *Educational Psychology Review, 15*(1), 1–40.

Di Martino, P., & Gregorio, F. (2018). The first-time phenomenon: Successful students' mathematical crisis in secondary-tertiary transition. In E. Bergqvist, M. Österholm, C. Granberg, & L. Sumpter (Eds.), *Proc. 42th Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 2, pp. 339–346). PME.

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. https://doi.org/10.1016/j.cedpsych.2020.101859

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

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. https://doi.org/10.1007/s11858-017-0893-1

Liebendörfer, M., & Göller, R. (2016). Abschreiben von Übungsblättern – Umrisse eines Verhaltens in mathematischen Lehrveranstaltungen [Copying homework tasks – analysis of a behaviour in mathematics courses]. In W. Paravicini & J. Schnieder (Eds.), *Hanse-Kolloquium zur Hochschuldidaktik der Mathematik 2014* (pp. 119–141). WTM.

Marsh, H. W., Pekrun, R., Parker, P. D., Murayama, K., Guo, J., Dicke, T., & Arens, A. K. (2019). The Murky Distinction Between Self-Concept and Self-Efficacy – Beware of Lurking Jingle-Jangle Fallacies. *Journal of Educational Psychology, 111*(2), 331–353. https://doi.org/10.25656/01:18125

Pajares, F. M., & Miller, D. M. (1994). Role of self-efficacy and self-concept beliefs in mathematical problem-solving: A path analysis. *Journal for Educational Psychology, 86*(2), 193–203.

Parker, P. D., Marsh, H. W., Ciarrocji, Marshall, S., & Abduljabbar, A. S. (2014). Juxtaposing math self-efficacy and self-concept as predictors of long-term achievement outcomes. *Educational Psychology: An International Journal of Experimental Educational Psychology, 34*(1), 29–48.

Rach, S. & Heinze, A. (2011). Studying Mathematics at the University: The Influence of Learning Strategies. In B. Ubuz (Eds.), *Proc. 35th Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 4, pp. 9–16). PME.

Rach, S., Heinze, A. & Ufer, S. (2014). Welche mathematischen Anforderungen erwarten Studierende im ersten Semester des Mathematikstudiums? [Which Mathematical Demands do Students Expect in the First Semester of Their Study?] *Journal für Mathematik-Didaktik, 35*(2), 205–228.

Rach, S., Kosiol, T., & Ufer, S. (2017). Interest and self-concept concerning two characters of mathematics: All the same, or different effects? In R. Göller, R. Biehler, R. Hochmuth & H.-G. Rück (Eds.), *Didactics of Mathematics in Higher Education as a Scientific Discipline – Conference Proceedings* (pp. 295-299). Universitätsbibliothek Kassel.

Rach, S., Ufer, S., & Kosiol, T. (2019). Self-concept in university mathematics courses. In U. T. Jankvist, M. Van den Heuvel-Panhuizen & M. Veldhuis (Eds.), *Proc. 11th Congress of the European Society for Research in Mathematics Education*. Freudenthal Group & Freudenthal Institute, Utrecht University and ERME.

Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary educational psychology, 61*, 101860. https://doi.org/10.1016/j.cedpsych.2020.101860

Schukajlow, S., Leiss, D., Pekrun, R., Blum, W., Müller, M., & Messner, R. (2012). Teaching methods for modelling problems and students' task-specific enjoyment, value, interest and self-efficacy expectations. *Educational Studies in Mathematics, 79*, 215–237.

Schukajlow, S., Rakoczy, K., & Pekrun, R. (2023). Emotions and motivation in mathematics education: Where we are today and where we need to go. *ZDM Mathematics Education, 55*, 249–267. https://doi.org/10.1007/s11858-022-01463-2

YOUTUBE CONTENT CREATORS' DISCOURSE: A MULTIPLE CASE STUDY ON THE CROSS PRODUCT USING COMMOGNITION AND POSITIONING THEORY

Farzad Radmehr, Kristin Krogh Arnesen, and Anita Valenta

Norwegian University of Science and Technology (NTNU)

Many university students turn to YouTube as a learning resource to reinforce their mathematical learning. However, there is a lack of research in mathematics education on the learning potentials of this type of resource. Through a multiple case study with two cases, we utilize commognition and positioning theory to investigate (a) what types of mathematical discourse are demonstrated and (b) how YouTube content creators position themselves and their viewers in the learning resources on the cross product. The findings indicate that different types of mathematical discourse are promoted (i.e., rituals vs. explorations), and different positioning occurs on the cross product (e.g., similar to many tutors helping students to get correct answers vs. promoting a storyline that mathematics makes sense, similar to discourse of many mathematicians).

INTRODUCTION

YouTube learning resources are among the most popular resources university students use to support their mathematical learning (Aguilar & Esparza Puga, 2020; Pepin & Kock, 2021). Past research suggests that students use YouTube learning resources to recall (Kanwal, 2020) and learn (Aguilar & Esparza Puga, 2020) mathematics. Although YouTube videos are widely used as learning resources, it seems that research in mathematics education has, so far, not sufficiently focused on the learning potential that emerges from these resources. Our study aims to contribute to this matter by taking a discursive approach, utilizing commognition and positioning theory.

Within commognition, mathematics can be seen as a particular discourse where narratives about mathematical objects are derived and some particular routines are regularly employed (Sfard, 2008). Mathematics learning takes place through participation in the mathematical discourse and its individualization, thus, learning is to become able to communicate in mathematics with oneself and others. We suggest that YouTube learning resources can be viewed as a context for learning and teaching, sharing similarities with classrooms where the content creators promote some particular ways of participating in mathematical discourse. Yet, YouTube learning resources also differ significantly from traditional classroom settings. Unless the content creator partakes in live streaming, direct interaction with students is absent.

In a classroom, the teacher can promote different kinds of participation in the discourse (Sfard, 2017), ideally bringing the students' discourse closer to the mathematicians' discourse. YouTube content creators are not necessarily teachers and do not need to act as such. Instead, the content creator can play the role of, for example, a friend

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. 4, pp. 41–48). PME.

assisting with exam preparations, or someone eager to share the beauty of mathematics. The observed variation of roles in YouTube learning resources led us to also adopt positioning theory (Harré et al., 2009) in our study. In the study, we seek to answer the following research questions:

1. *What types of mathematical discourse are demonstrated by YouTube content creators in mathematical learning resources on the cross product?*

2. *How do YouTube content creators position themselves and their viewers in mathematical learning resources on the cross product?*

To answer the research questions, we analyse two popular YouTube learning resources on the cross product. We chose this mathematical object because it is typically included in introductory linear algebra courses, and it has a number of applications in mathematics and other disciplines.

THEORETICAL BACKGROUND

We use commognition (Sfard, 2008) as the main theoretical framework. Commognition is a discursive theory, where learning mathematics is conceptualised as changes in participation in a mathematical discourse. In addition, we use positioning theory (see Harré et al., 2009) to capture nuances in how the YouTube content creator participates in, and thus shapes, the discourse. The two theories both centre around discursive and social aspects and have previously been combined to study students' identity in relation to learning mathematics (e.g., Heyd-Metzuyanim & Cooper, 2022).

Commognition

In commognition, mathematical discourse consists of four elements: The words that have specialised meaning within the discourse, the visual mediators used to represent mathematical objects, the narratives that define and describe mathematical objects and their properties, and the routines—the actions that are regularly performed within the discourse (Sfard, 2008). Sfard (2008) describes three types of routines: The *rituals* are performed with the social aim of "fitting into" the discourse (e.g. when students apply a known algorithm because they are expected to do it). In contrast, the *deeds* and the *explorations* are performed to change the discursive objects (e.g. when you add two numbers as part of a problem-solving process), or to create or substantiate the narratives about them (e.g. when students investigate patterns, searching for a conjecture), respectively. Explorative participation is the aim of learning mathematics, but rituals are seen as a necessary part of the trajectory (Sfard, 2008).

Sfard's notions of discourse and routines have been extended to essentially any human activity within communities (Lavie et al., 2019). In the mathematics classroom, there is a certain mathematical discourse. However, because the teacher is usually more experienced in "canonical" mathematical discourse than the students, and because the teacher's aims are different from those of a mathematician, the classroom discourse is not exclusively mathematical (even if we exclude the social or upbringing parts of the discourse). The same is true for YouTube learning resources. In this study, we use positioning theory to add another layer when investigating the discourse in the YouTube learning resources.

Positioning theory in the context of YouTube learning resources

Originating in social psychology, positioning theory aims to describe how people interact with an emphasis on the role of rights and duties within a discourse (Harré et al., 2009), and is thus closely related to concepts such as agency, power, and authority. The theory has been adapted to educational research, where it provides a lens to "the in and over time construction of positioning actions of teachers and students in developing episodes for learning and participating in classrooms" (Green et al., 2020, p. 119). Two fundamental terms of positioning theory are *positioning* and *storylines*. Positioning happens through acts where people are attributed to, or allow themselves, rights and duties. Examples are the positions of "teacher" (with a certain authority in the classroom, and a strong duty of scaffolding students' learning) or "mathematician" (with an indisputable mathematical authority). Positioning is situated in the context of storylines, "social and discursive practices within which people are embedded that inform actions" (Green et al., 2020, p. 121), that again are placed within social, historical and cultural situations (Harré et al., 2009). Thus, storylines can be well-known histories like "David and Goliath" or "Mathematics is a boring and difficult school subject", but also ad hoc constructed stories.

Earlier commognitive research has shown how positioning acts and storylines are related to the type of discourse that is provided by the teacher, as well as influencing students' opportunities to learn. Although neither uses the term "positioning", Sfard (2017) and Heyd-Metzuyanim et al. (2015) describe teachers who repeatedly position themselves and their students and evoke certain storylines about mathematics and the learning thereof. The lesson learnt from these studies is that the teacher's positioning strongly affects the mathematical discourse. In both cases, the teacher's positioning reinforce ritual participation and thus limit the students' opportunities to participate in an explorative discourse. These studies focused on the teacher's interactions with students. In a YouTube learning resource, such communication is limited. Still, it is possible to recognise storylines, for example, by looking at how mathematics is alleged to and justified.

METHODOLOGY

The study reported here is a multiple case study with two cases. To demonstrate that discourse is independent of the mathematical topic in question, we examine two YouTube learning resources on the same mathematical topic, cross product, in this study. To find YouTube learning resources, we searched YouTube using the keyword "cross product" in October 2023, and sorted the results based on view counts. The first two learning resources in English were chosen as the cases for this study. The first case[1] is from the *3Blue1Brown* channel. The video is viewed 1.6 million times as of 19th December 2023. The second case is from *The Organic Chemistry Tutor* channel, which has 7.2 million subscribers. The video on the cross product in this channel[2] has been viewed 1.2 million times as of 19th December 2023.

Data analysis

The analysis was carried out in two parts, corresponding to the two research questions. In the first part, we looked mainly for evidence of explorative or ritual participation in the mathematical discourse in the two cases. This was done by a deductive analysis of the content creator's routines: Explorations are product-oriented, with a focus on "constructing and endorsing a new narrative about mathematical objects" (Lavie et al., 2019, p. 166), whilst rituals typically focus on "manipulation of mathematical symbols, without any reference to the objects signified by them" (Heyd-Metzuyanim et al., 2015, p. 548). Moreover, rituals and explorations can be distinguished by looking at where the action resides: Referring to the actions undertaken by participants ("do this") is related to rituals; explorations, on the other hand, are removed from the humans performing them, concentrating on the mathematical objects (Heyd-Metzuyanim et al., 2015).

In the second part, we identified routines that revealed the content creator's positioning of himself or the viewers, and the mathematics-related storylines that were alluded to or explicitly invoked. The former involves discourse that reveals the content creator's aims, relation to mathematics, relation to the audience and so on, as well as talk about who the audience is, or assumptions of what they want. The latter involves talk about the nature and goals of mathematics, and about how to learn mathematics. This part of analysis was data-driven because positions or storylines were not predetermined.

RESULTS

Case 1

The content creator introduces the video by saying:

> Last video, I talked about the dot product, showing both the standard introduction to the topic, as well as a deeper view of how it relates to linear transformations. I'd like to do the same thing for cross products, which also have a standard introduction along with a deeper understanding in the light of linear transformations.

By pointing to "deeper understanding" and relations between mathematical objects, the content creator is indicating that the discourse he invites the viewers into is about mathematical objects, their properties and relations, and thus explorative. The following excerpts strengthen this claim (time into the video indicated in parenthesis). The content creator introduces a definition of cross product in 2D (sic) by relating it to the parallelogram spanned by them:

> (1) The cross product of $\bar{v}$ and $\bar{w}$, written with the X-shaped multiplication symbol, is the area of this parallelogram. (1:09)

Here, he talks about the objects and gives a definition, before he further presents a procedure to find such product:

(2) For the 2-D cross-product $\bar{v} \times \bar{w}$, what you do is you write the coordinates of $\bar{v}$ as the first column of the matrix, and you take the coordinates of $\bar{w}$ and make them the second column then you just compute the determinant. This is because a matrix whose columns represent $\bar{v}$ and $\bar{w}$ corresponds with a linear transformation that moves the basis vectors $\hat{\imath}$ and $\hat{\jmath}$ to $\bar{v}$ and $\bar{w}$. The determinant is all about measuring how areas change due to a transformation. And the prototypical area that we look at is the unit square resting on $\hat{\imath}$ and $\hat{\jmath}$. (3:13)

Even though the content creator's utterance here is about a procedure, the focus is on the product (thus, it is a deed and not a ritual), and he also tries to justify the process by relating it to linear transformations. The content creator sums up the presentation on 2D-vectors by saying:

(3) As with any new operation you learn, I'd recommend playing around with this notion just to get kind of an intuitive feel for what the cross product is all about. (4:17)

Later, he points out that the operation defined in 2D is not actually a cross product, and he proceeds to a definition of cross product in 3D by defining the length of the new vector, and its orientation. As before, the content creator's utterance is about mathematical objects and the construction of narratives about them, as he presents the process for finding a cross product of two vectors. Here, he does not spend time on calculating the determinant, he simply refers to a previous video in case the viewer does not remember how to proceed. Rather, he points out:

(4) Now, this process looks truly strange at first. You write down a 3D matrix where the second and third columns contain the coordinates of $\bar{v}$ and $\bar{w}$. But for that first column you write the basis vectors $\hat{\imath}$, $\hat{\jmath}$ and $\hat{k}$. Then you compute the determinant of this matrix. The silliness is probably clear here. What on earth does it mean to put in a vector as the entry of a matrix? (7:24)

The creator proceeds by trying to give meaning to the procedure, pointing out that it yields a linear combination of the basis vectors $\hat{\imath}$, $\hat{\jmath}$ and $\hat{k}$. Still, he says that "students are told just to believe it" that what one gets by performing the procedure is the unique vector which was defined earlier as the cross product. He remarks that "It's not just a coincidence", that it is about the idea of duality and that the argument is somehow complex but is provided in another video for those who are interested. At the end, he stresses that it is important to know what the cross product represents geometrically.

As illustrated by the excerpts, the content creator of this video employs explorative mathematical routines—the focus is always on mathematical objects (vectors, cross products, linear transformations, determinants, parallelograms, and areas), and the construction and substantiation of new narratives about their properties and relations.

Regarding the content creator's positioning, he promotes a storyline that mathematics makes sense, that it is about mathematical objects and relations between them, about constructing new objects (as cross product), narratives about how they are related to other objects (as determinants, parallelograms, 2D and 3D) and justifying such relations: even though he does not go into why the procedure gives the vector which is defined as the cross product, he indicates that there *is* an explanation. Furthermore, the creator promotes a storyline that learning mathematics involves "playing with new operations" (as in excerpt (3)), sense-making and asking *why* (as in excerpt (4)). We suggest that the creator positions himself as a mathematician in the video, giving an example of how objects are constructed (e.g. transition from 2D to 3D), and how mathematicians proceed to explore mathematical objects (particularly seen in excerpts (3) and (4)). Also, he positions the students as sense-makers, by remarking that they often are asked to "just believe it", but that there is an explanation which he makes available for them even though it is complex.

Case 2

The content creator of this video starts by pointing out that "In this video we're going to talk about how to find the cross product of two vectors", which indicates that a procedure will be in focus. And it is, as shown in the excerpts below. He starts with two arbitrary vectors, "Let's a vector a is $3i$ plus $5j$ minus $7k$, and vector b is $2i$ minus $6j$ plus $4k$" and continues:

> (1) So, what is the cross product of vectors a and b, how can we find the answer. What I'd like to do is first put this in the form of a matrix this is $i\ j\ k$ which corresponds to the x, y and z components of a vector and then first we need to put vector a in the middle (…) And basically, you got to find the determinant of this 3×3 matrix. (1:21)

The content creator continues with step-by-step calculations, as "12 minus negative 14 is like 12 plus 14 that's 26 and we have a minus in front so it's going to be negative 26". Even though the content creator starts with the phrase "what is the cross product of vectors a and b", his focus is completely on "doings"—what a person needs to do to find the cross product of the two vectors, not what the cross product actually is. After performing the calculation, and getting a vector c as a result, he suggests that one can check whether the result is correct:

> (2) If you take the cross product of those two vectors, you're going to get another vector, vector c that's perpendicular to a and b. (4:31)

He continues by reminding the viewers of how to check whether two vectors are perpendicular (is the dot product 0?). Then, he calculates the dot product of the vectors c and a, and c and b, which both are 0, and he concludes that the cross-product calculation was correct. Here, the creator presents a property of the cross product (which was used as a definition of cross product in Case 1) as a way to check the calculation. Even though he talks about a property of a mathematical object, the focus is on a person doing a procedure, and not on the mathematical object. Thus, the creator's routines are rituals.

Next, the creator gives a new example of vectors a and b and performs the same procedure, step-by-step, again pointing out at the end that, "just to make sure that we have the right value", one can check the dot-products, and he does that in the video.

As already pointed out, the mathematical routines the content creator engages in, and invites the viewers into, are rituals. He never talks about what mathematical objects are or which properties they have, the focus is on what to do to find a cross-product. The property of the cross product, i.e., that it is a vector perpendicular to the factor vectors, is mentioned, but only in the context of a procedure to verify correctness.

Concerning positioning, the content creator presents mathematics as performing specific procedures, step-by-step, correctly. There is no flexibility and no justifications or sense-making ("you got to …", excerpt (1)). He never mentions students or learning mathematics explicitly, but his detailing of the procedure and emphasis on checking the results indicates that the storyline about learning mathematics is about being able to perform some given procedures correctly. We suggest that he is positioning himself as a tutor, helping students to get correct answers on tasks asking to find a cross product, and that he is positioning learners as interested in only that.

DISCUSSION

In this study, we employ commognition and positioning theory to analyse the discourse in two YouTube learning resources. There is a lack of research on the learning potential of YouTube resources, and our study aims to address this gap. Furthermore, our utilization of a combination of commognition and positioning theory is novel, although one could argue that positioning was implicitly used in earlier commognition studies earlier (e.g. Heyd-Metzuyanim et al., 2016; Sfard, 2017).

As the analysis shows, the two cases are very different. While the first demonstrates a truly explorative mathematical discourse, the second is an example of a discourse dominated by rituals. Naturally, an important aspect in the learning of the cross product is the procedure for how to find it, and introducing some mathematical procedure can easily become process-oriented, like in Case 2. In Case 1, however, the creator is always focused on the result of the process, and not the process itself (excerpt (2)): He is not dwelling on technicalities, but rather focuses on the sense-making and justification of the procedure. In both cases, the analysis shows a clear connection between the mathematical discourse and the positionings undertaken by the content creators. The storylines about mathematics were mostly implicitly presented in both videos, aligning well with the characteristics of the respective discourses.

In the classroom, students often have little or no choice but to be there, and thus the kind of discourse orchestrated by the teacher is crucial for students' opportunities to learn (Heyd-Metzuyanim et al., 2015). On YouTube, students can choose what they want to watch, depending on what they want to learn. However, the commognitive discourse analysis performed above reveals important characteristics of the mathematical learning opportunities available on YouTube. By adding positioning to the analysis, we revealed positionings and storylines about mathematics that also affect the potential learning opportunities. Rituals, that dominate in Case 2, are important for learning (Sfard, 2008). However, in the same video, there is no mention of "something more" than procedures: the storyline is about mathematics as procedures and *only that*, and this can limit further learning opportunities.

We propose that combining positioning theory with commognition effectively describes the discourse set up in YouTube learning resources. positioning may be more flexible and visible in YouTube learning resources compared to traditional mathematics teaching, we suggest that employing positioning theory along with commognition can enhance discourse analysis in both contexts.

Notes

1. Cross products | Chapter 10, Essence of linear algebra (youtube.com)
2. Cross Product of Two Vectors Explained! (youtube.com)

References

Aguilar, M. S., & Esparza Puga, D. S. E. (2020). Mathematical help-seeking: observing how undergraduate students use the Internet to cope with a mathematical task. *ZDM*, *52*(5), 1003–1016. https://doi.org/10.1007/s11858-019-01120-1

Harré, R., Moghaddam, F. M., Cairnie, T. P., Rothbart, D., & Sabat, S. R. (2009). Recent advances in positioning theory. *Theory & psychology, 19*(1), 5–31. https://doi.org/10.1177/0959354308101417

Heyd-Metzuyanim, E., Tabach, M. & Nachlieli, T. (2016). Opportunities for learning given to prospective mathematics teachers: between ritual and explorative instruction. *Journal of Mathematics Teacher Education 19*, 547–57. https://doi.org/10.1007/s10857-015-9311-1

Heyd-Metzuyanim, E., & Cooper, J. (2022). When the Problem Seems Answerable yet the Solution is Unavailable: Affective Reactions Around an Impasse in Mathematical Discourse. *International Journal of Research in Undergraduate Mathemathics Education*. https://doi.org/10.1007/s40753-022-00172-1

Green, J., Brock, C., Baker, W. D., & Harris, P. (2020). Positioning theory and discourse analysis. In N. S. Nasir, C. D. Lee, R. Pea, & M. McKinney de Royston (Eds.), *Handbook of the cultural foundations of learning* (pp. 119–140). Routledge.

Kanwal, S. (2020). Exploring affordances of an online environment: A case-study of electronics engineering undergraduate students' activity in mathematics. *International Journal of Research in Undergraduate Mathematics Education, 6*, 42–64. https://doi.org/10.1007/s40753-019-00100-w

Pepin, B., & Kock, Z. J. (2021). Students' use of resources in a challenge-based learning context involving mathematics. *International Journal of Research in Undergraduate Mathematics Education*, *7*(2), 306–327. https://doi.org/10.1007/s40753-021-00136-x

Sfard, A. (2008). *Thinking as communicating: Human development, the growth of discourses, and mathematizing.* Cambridge University Press.

Sfard, A. (2017). Ritual for ritual, exploration for exploration. In J. Adler, & A. Sfard (Eds.), *Research for educational change: Transforming researchers' insights into improvement in mathematics teaching and learning* (pp. 39–63). Routledge.

FROM INNOVATION TO IMPACT: FACTORS SHAPING THE SCALING SUCCESS OF THE TRIUMPHS PROJECT

Iresha Ratnayake[1], Linda Marie Ahl[1], Johan Prytz [1], and Uffe Thomas Jankvist[1,2]

[1]Uppsala University, Sweden, [2]Aarhus University, Denmark

The paper recounts the successful implementation story of the TRansforming Instruction in Undergraduate Mathematics via Primary Historical Sources (TRIUMPHS) project. Our analysis of the project involved examining influential factors (Century & Cassata, 2016) and scaling dimensions (Coburn, 2003). We identified how influential factors and strategies employed by the TRIUMPHS project positively impacted scalability, particularly highlighting sustainability. These findings underscore the importance of innovation, user engagement, and the operational context in driving project expansion and long-term viability.

INTRODUCTION

In implementing innovation, especially in discussions about scaling, leaders of reform aspire for the innovation to have a widespread impact on the users (Aguilar et al., 2023). Scaling the heights of successful implementation projects requires a nuanced understanding of the underlying factors contributing to their triumphs. This paper narrates the success story of an implementation project—TRansforming Instruction in Undergraduate Mathematics via Primary Historical Sources (TRIUMPHS)—focusing on the factors that enabled the scaling of the project. We employed the factors of influence proposed by Century and Cassata (2016) and the scaling dimensions proposed by Coburn as the theoretical constructs to address the research question: *What factors influenced the success of scaling of the TRIUMPHS project?*

Following a comprehensive explanation of the theoretical constructs, we provide detailed descriptions of the TRIUMPHS project. After the methods and the results sections, the paper concludes with a discussion on influential factors, the projects' sustainability, and a final summarising conclusion.

THEORETICAL CONSTRUCTS

Exploring implementation research serves as a driving force to unearth the determinants influencing the effectiveness of adopting innovations, spotlighted by Century and Cassata (2016). These determinants, labelled as factors of influence, encapsulate characteristics tied to four distinct realms: the innovation itself, the users, the organizational setting, and the external environment (Century & Cassata, 2016). The specific attributes defining these factors of influence are contingent on the particular context under scrutiny, spanning diverse domains like healthcare, sports organizations, manufacturing, and, notably, educational institutions—our focal point herein, primarily concentrating on tertiary educational institutes.

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. 4, pp. 49–56). PME.

In line with Century and Cassata (2016), we define innovations as the pivotal components targeted for alteration or enhancement, particularly those emerging from research in mathematics education. Key intrinsic factors of influence concerning the innovation involve its adaptability, relevance to end-users, and alignment with specific operational practices. Educators, chiefly teachers or instructors, are the primary users wielding significant influence as change catalysts within educational settings, including universities. Factors shaping implementation projects involving educators encompass their grasp of innovation, proficiency in mathematical and pedagogical domains, prior experiences, organizational adeptness, classroom management style, and an array of affective traits like beliefs, values, attitudes, motivation, self-efficacy, and openness to novel approaches (Century & Cassata, 2016).

Organizational factors of influence stem from decisions made by stakeholders within educational institutions, spanning choices related to class sizes, resource allocation, physical infrastructure, scheduling, and the overarching organizational framework governing instructors' endeavours. Administrative procedures, management methodologies, and policy determinations linked to the specific innovation considered constitute aspects of these organizational factors. Meanwhile, external environmental factors comprise opportunities and constraints lying beyond the control of educational institution stakeholders, influencing the implementation of innovations. These external influences typically transcend the school's sphere of influence and can encompass economic conditions, infrastructure limitations, and shifts in political priorities. Tangible examples illustrate external influences, such as inadequate network connectivity impeding instructor support for professional development or alterations in budget allocation due to political leadership shifts.

Delving into the dissemination of innovations, Coburn (2003) presents four interconnected dimensions of scale: depth, sustainability, spread, and shift in reform ownership. Depth refers to a change in classroom practices surpassing mere shifts in curriculum resources or the introduction of specific teaching methods and activities. Coburn (2003) argues that scaling in depth includes altering teachers' beliefs and norms regarding communication and pedagogical practices, requiring a fundamental change in ideas regarding effective instruction and student learning for successful implementation.

For implementation projects, it is essential to devise a strategy ensuring the innovation's continuity within the organization beyond the project's completion. Coburn introduces the dimension of sustainability within scaling, emphasizing the support structure required to uphold the innovation's vitality even after the withdrawal of support from reform leaders (Coburn, 2003).

While expanding the innovation to other educational settings remains a crucial aspect of scaling, Coburn (2003) delineates the concept of *spread* across various levels of stakeholders within the school system: classrooms, schools, and districts. The fourth dimension of scaling is a shift in reform ownership as part of the scaling process

(Coburn, 2003). Initially, the genesis of an innovation lies with its creators and implementation architects. However, for the innovation to seamlessly integrate into the organization, the authority to scale its implementation must gradually transition to districts, schools, and teachers. This shift allows for scaling in-depth, sustainability, and spread to maintain effectively over time.

THE TRIUMPHS PROJECT

The TRIUMPHS project, funded by the National Science Foundation (NSF), represents an American initiative that leverages prior experiences using primary historical sources from the history of mathematics to transform undergraduate mathematics instruction (Clark et al., 2022). This endeavor centers on primary source projects (PSPs), a unique approach intertwining historical materials within curricular resources and assessing the anticipated educational advantages. On the one hand, TRIUMPHS is an innovative development project. On the other hand, it is a research initiative exploring students' mathematical understanding, influenced by Sfard's (2008) theory of commognition.

Primary historical sources in this project encompass excerpts from manuscripts, letters, and other works by past mathematicians. These sources delve into topics like the evolution of mathematical concepts, such as Leonhard Euler's (1707-1783) contributions or the abstract algebra development involving mathematicians like J. L. Lagrange (1736–1813), Augustin Cauchy (1789–1857), and Arthur Cayley (1821–1895), as previously explored in TRIUMPHS-related projects. The PSPs present segments of selected primary sources, often with English translations through a 'guided reading' approach. These readings provide context, clarify unfamiliar terminologies, and offer tasks to enhance students' content comprehension.

These PSPs are designed for undergraduate students, target college and university instructors and are envisioned as alternatives to conventional textbooks or lecture materials. Spanning the same duration as traditional coursework, they encompass self-contained mathematical and historical content. Notably, PSPs offer instructors the flexibility to adapt the materials to suit their needs, considering them as no-cost, open-source LaTeX files. Mathematics-related degree holders, i.e., college and university instructors, are invited to specialized training seminars where their creators introduce PSPs. These seminars also feature 'site-testers' who share their experiences using different PSPs. Over time, experienced instructors may even develop their own PSPs under project guidance.

The TRIUMPHS project leaders aim for 'implementation at scale', seeking to expand the innovation to broader contexts encompassing more undergraduate students, instructors, colleges, universities, and districts (Barnett et al., 2022). Barnett et al. (2022) discuss various factors related to PSP attributes, dissemination, and implementation support strategies. In terms of impact, Barnett et al. (2022) cite positive experiences, both quantitatively and qualitatively.

METHODS

We conducted semi-structured interviews—each 90 minutes long—via Zoom with five out of seven principal investigators (PIs) of the TRIUMPHS project in September and October 2023. The interview protocol comprised six themes with probing questions to ensure comprehensive information. The deliberate use of open-ended questions aimed to elicit extensive insights. The six themes were background, the role of mathematics education theories, materials (PSPSs), implementation, factors influencing, and evaluation. The interviews were transcribed using Zoom's transcription feature and manually reviewed for clarity. For this paper, we analyzed the responses for the last four themes: materials (PSPSs), implementation, factors influenced, and evaluation. Two theoretical frameworks, influential factors of implementation research (Century & Cassata, 2016) and Coburn's (2003) scaling dimensions, were incorporated in analyzing the transcripts. Additionally, we referenced the project's final report alongside the interview transcripts to gather precise data figures. For anonymity, we used the pseudonyms PI# for PIs.

RESULTS

This section presents the results of the analysis of the interviews. We structure them under Coburn's (2003) four dimensions of scaling: depth, spread, shift in reform ownership, and sustainability, respectively.

Depth: Initially, PSPs were written exclusively for a limited number of courses, such as discrete mathematics. Over time, many PSPs were designed to encompass all undergraduate math courses. as noted by PI4.

> PI4: Before it was mostly like discrete math, the primary source projects were done in [...]. But we wanted to be able to write projects in pre-calculus, abstract algebra, linear algebra, geometry, topology, and real analysis. The entire gamut of courses. So that no matter what course you were teaching in an undergraduate math curriculum, you would find at least one primary source project.

All five PIs we interviewed confirmed PI4's assertion. They identified several reasons that could motivate authors to compose PSPs for various courses. For instance, if an author comes across a captivating primary source, they might create a PSP related to a topic taught in their classroom. Similarly, suppose an instructor is dissatisfied with their teaching approach and considers experimenting with a new method. In that case, they may seek a primary source linked to that subject and develop a PSP. The PIs emphasize the importance of authors connecting historical sources to new mathematical topics. For instance, PI1 articulated, "We were always hoping that whatever the author picked for a source, they could tie it to a standard topic."

Spread:

The PIs contacted instructors, inviting them to act as site testers through various channels, including platforms like the Mathematical Association of America, email correspondence, daily newsletter subscriptions, workshops, and more. They aimed to have a PSP tested by at least one site tester other than the author.

> PI1: Some of our projects were very widely tested. And I think we ended up with, out of a hundred projects, about ten that were only tested by their author and not by another individual. Maybe less than ten.

However, a few PSPs underwent testing solely by their respective authors. Nevertheless, PI1 and PI2 verified that there were fewer than five dropouts among site testers, primarily attributed to personal reasons such as changes in employment. The PIs stressed that favorable student feedback during evaluations significantly impacted instructors' decisions to utilize a PSP repeatedly.

> PI4: I think a lot of the site testers became repeat users. I wouldn't say all, but a lot. In fact, I would guess a majority

Shift in reform ownership: The PSPs were freely available and were open to modifications.

> PI5: These materials were available to a wide range of professors, instructors, [and] teachers. They were free open source, and then, unlike textbooks, you can edit [...] these. You can change everything. You have the [LaTeX] code, for example.

He highlighted that, apart from granting instructors agency as users, they also had the opportunity to receive support from the PSP authors. PI3 echoed PI5's statement.

> PI3: Because it was important that we made these projects publicly available, people could grab them without having to ask permission to do so. They were invited to contact the authors of the projects to recover the raw LaTeX code. So that [the] users could modify them as needed for their actual implementation.

This empowerment enabled instructors to choose materials (PSPs), modify them, and implement them in a way they wanted to adapt them according to their preferences. Subsequently, some instructors expressed interest in creating new PSPs and received continuous support until the PSPs met the required standards. PI1, in particular, provided one-on-one support to these authors.

> PI1: Every project was read by at least three people and got two stages of review. I worked a lot with authors at the beginning and at the end of polishing: [From] launching a project, picking sources, and then the final polishing of them.

Further, a statement from PI2 confirms a change in reform ownership.

> PI2: Higher ed[ucation] implementation is significantly different from K-12. And I think because of that, it felt like a serendipitous kind of outcome like the folks that were all dialed in and who are now our external authors.

Hence, over time, ownership was transferred to the instructors.

Sustainability: Addressing the sustainability of innovation, we asked for information about the future of TRIUMPHS.

Interviewer: What is next? How could the results from TRIUMPHS be spread and implemented in the future?

PI1: We have already formed a new TRIUMPHS society. So, towards the end of the grant, we started thinking seriously about how we can recruit the next generation of people who will keep this [...] initiative going. [...] It will probably focus more on the development and implementation side. There is a peer-reviewed journal that will be associated with that effort.

Moreover, even after the project's funding concluded and officially ended in July 2023, the community has remained dynamic, actively authoring and uploading PSPs.

PI4: we are still coming up with newer and better versions of the PSPs. There are still new PSPs coming out and being posted [on the website. The new PSPs indicate that the newly formed society has remained engaged in developing and modifying existing PSPs based on the evaluations.

DISCUSSION

In this section, we sought to provide answers to the research question. In line with influential factors of implementation research pointed out by Century and Cassata (2016), innovation likely influenced the scale of the TRIUMPHS project. The apparent *innovation* of the TRIUMPHS project is the PSPs. Using primary sources to learn mathematical concepts was a novel experience for the students. More importantly, understanding the language and historical mathematics was undoubtedly a challenge. To overcome this challenge, the project team carefully structured the PSPs in a specific manner. For example, each PSP revolves around one or several primary source texts. The tasks aid students in comprehending and analysing the primary sources. They offer unfiltered insights, expressed in language unaltered by decades of refinement. The context in which the authors write, often involved in evident discoveries, adds to the appeal for student engagement. The context in which authors write engages students, and the tasks help students decipher the text and understand the mathematical content, linking it to modern curricula. In that way, the *innovation* is relevant to the end users and aligned with the standard practice of learning new mathematical concepts.

In addition to the structure of the PSPs, the innovation also paid attention to instructors as users. To make the instructors (as *users*) comfortable teaching with the PSPs, a set of detailed instructions encompassing guidance on incorporating tasks into classroom settings and offering insights on their utilization was integrated within the implementation schedule of each PSP. This development stands out as another significant contribution of the TRIUMPHS project, particularly in providing a valuable resource for instructors.

Another vital aspect of the innovation is its adaptability. The flexibility allows instructors to tailor and adapt the PSPs to their needs. This freedom empowered instructors to create their materials for diverse mathematics courses. That was a reason for the *depth* (Coburn, 2003) of the TRIUMPHS project—expansion to almost all undergraduate mathematics courses; by the end of the project in July 2023, 99 PSPs had been developed (Clark & Barnett, 2023).

The primary end users of the project are tertiary-level mathematics instructors. The project team has organized numerous workshops for groups of high school, college, and university teachers who volunteered as site testers. Site testing is entirely voluntary, and a key factor for instructors to remain engaged in the project is their satisfaction with using PSPs. Consequently, within the initial six years of the grant, "133 instructors officially used PSPs, … for a total of 436 distinct implementations in over 240 classrooms at 109 different institutions across the US and Canada" (Clark & Barnett, 2023, p.3). Hence, TRIUMPHS gradually *spread*.

Over time, some site testers became authors of PSPs, receiving ongoing support from the project team. They were guided through drafting, receiving feedback from a PI, and modifying PSPs until achieving a high-quality standard. As the new authors gained proficiency, they transitioned from modifying existing resources to creating entirely new ones. Gradually, as users became more accustomed to PSPs and their role in contemporary mathematics education, instructors expanded the scope and took ownership of PSPs. This shift in ownership aligns with Coburn's (2003) scaling dimensions regarding reform ownership. However, the agency given to instructors might be attributed partly to these individuals being university instructors. They are experts in the field of their teaching. This scenario might differ considerably from K-12 implementation, where such autonomy is not typically observed.

The TRIUMPHS project most likely had a *sustainability* plan to ensure the continuity of innovation beyond the project conclusion. As the project progressed, the PIs and site testers— comprising authors and exclusive users —evolved into a community of practitioners. The PIs of the TRIUMPHS project announced this community as the TRIUMPHS Society. Even after the end of the project, the TRIUMPHS society has been active in writing new PSPs and upgrading existing ones based on evaluations.

Organizational influences appeared favourable for the users in their settings. The PIs did not complain about the support the users and authors received from the institutes or the infrastructure facilities. There were no external influences on the project due to the autonomy of tertiary-level instructors in deciding what to teach, how to teach, and how to evaluate. Not surprisingly, in organizational settings, tertiary-level institutes in developed countries appeared to receive preferential support for such projects. However, this may not be the case in a different context, for example, in a developing country or a K-12 setting.

CONCLUSION

This paper delineates influential factors contributing to the success of an implementation project in scaling. The apparent finding is that the TRIUMPHS project scaled up in terms of Coburn's (2003) scaling dimensions: depth, sustainability, spread, and shift in reform ownership. We identified the interplay of influential factors and strategies employed by the TRIUMPHS project, positively impacting project scalability. In particular, the innovation, including its distinctive structure and supplementary materials such as instructor guidelines, alongside its adaptability and the continual support offered by the authors, markedly influenced the TRIUMPHS project's success. Further, as Aguilar et al. (2023) pointed out, the well-developed relationship between producers (authors) and users, evolving into a collaborative co-production, will likely positively impact the TRIUMPHS project's success.

The findings of the paper confirm two of the three factors proposed by Aguilar et al. (2023)—contact factor (producer-user contact – a strong relationship between the authors and users in the TRIUMPH project) and material factor (e.g., incorporating detailed guidelines to the users)—are crucial factors in scaling up an implementation project. In addition to these factors, the adaptability of the materials and the plans protect the established bond between the producers and users (TRIUMPH society), which is also worth considering in future implementation projects.

ACKNOWLEDGEMENT

This paper is part of the 2020-04090 grant under the *Swedish Research Council*.

References

Aguilar, M. S., Ahl, L. M., Jankvist, U. T., & Helenius, O. (2023). Towards characterization of scale and scaling in implementation research within mathematics education. *Implementation and Replication Studies in Mathematics Education, 3*(1), 1-24.

Barnett, J. H., Clark, K. M., & Can, C. (2022). Transforming mathematics instruction via primary historical sources: A study of influential factors on implementation of a curricular innovation at the tertiary level. *Implementation and Replication Studies in Mathematics Education, 2*(2), 208–240.

Century, J., & Cassata, A. (2016). Implementation Research: Finding Common Ground on What, How, Why, Where, and Who. *Review of Research in Education, 40*(1), 169–215.

Clark, K. M., & Barnett, J. H. (2023). *TRIUMPHS Evaluation Report*. (Unpublished).

Clark, K. M., Can, C., Barnett, J. H., Watford, M., & Rubis, O. M. (2022). Tales of research initiatives on university-level mathematics and primary historical sources. *ZDM – Mathematics Education, 54*(7), 1507–1520.

Coburn, C. E. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher, 32*(6), 3–12.

Sfard, A. (2008). *Thinking as communicating.* Cambridge University Press.

TEACHER AGENCY AND THE USE OF CURRICULUM MATERIALS ACROSS CULTURAL CONTEXTS

Janine Remillard[a], Lara Condon[a], Tuula Koljonen[b], Heidi Krzywacki[c], and Riku Sayuj[a]

University of Pennsylvania[a], Linköping University[b], University of Helsinki[c]

Using an ecological and dynamic view of teacher agency, this study explores the relationships between teachers' professional actions and decisions, mathematics curriculum materials (CMs), and cultural norms and values at play in four educational contexts: Finland, Flanders (Belgium), Sweden, and the United States. The data were drawn from a survey of 397 teachers (grades 1-6), inquiring into self-reported use and perceptions of their CMs. Analysis of the most commonly reported CMs illustrated characteristics that reflect cultural values in each context. Survey findings indicated that teachers in all contexts use CMs purposefully and in relation to their own ideas about teaching. We also found context-specific differences in how teachers relied on CMs for different curricular aims, adding complexity to notions of CM use.

INTRODUCTION

Curriculum materials (CMs)[1] have long been a mainstay in mathematics classrooms around the world (Valverde et al., 2002). In many school systems, they are used to communicate curriculum policy and ensure coherence across classrooms. They are also seen as valuable classroom resources for teachers, supporting instructional planning and decision-making (Remillard, 2005). The roles CMs play in educational practice are influenced by educational policies, cultural norms and values, and teaching practices, which differ across cultural contexts (Pepin & Haggarty, 2001). Cross-cultural studies of mathematics teaching offer insights into the complex web of factors that contribute to classroom instruction in any cultural context (e.g., Stigler & Hiebert, 1999). At the same time, many studies highlight the role of textbooks as shaping forces and downplay the teacher's role in this process (e.g., Valverde et al., 2002).

This study explores the relationships between teachers' professional actions and decisions, mathematics CMs, and cultural norms and values in four educational contexts: Finland, Flanders (Belgium), Sweden, and the United States. We draw on a professional space lens (Oolbekkink-Marchand et al., 2017), which conceptualizes teacher agency as a dynamic process of negotiating and acting within objective and subjective features of an educational landscape. *By looking comparatively at how teachers in the four contexts describe their uses of mathematics CMs in relation to how typical CMs guide teaching actions, we seek to uncover context-specific differences in the interplay between teachers' professional space and the enactment of agency.*

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. 4, pp. 57–64). PME.

PROFESSIONAL SPACE AND THE ENACTMENT OF AGENCY

This study takes an ecological approach to teacher agency, viewing teachers as active agents in a system, navigating the tension between "their work as professional practitioners in the classroom, and their dependence on organizational structures, such as school and curriculum provided by state governance" (Wermke & Höstfält, 2014, p. 60). Oolbekkink-Marchand et al. (2017) use *professional space* to refer to objective components of school systems, along with teachers' subjective perceptions and negotiation of them. Rather than viewing teacher agency as objectively determined by context, professional space views teachers as "active interpreters of the school context and the space they have, to act on their own personal goals" (p. 38). Further, teachers act "by means of their environment," and agency is understood as an interplay among "individual efforts, available resources, and contextual and structural factors" (p. 38).

We adapt Priestley and colleagues' (2015) ecological model of teacher agency to guide our analysis of the interplay between CMs, specifically teacher's guides, and teacher decision-making in relation to their instructional goals (Fig. 1). Their model proposes that teachers enact agency in relation to influential components of their current context, which they refer to as the *practical-evaluative* dimension. This dimension includes cultural (values, beliefs), structural (policies), and material components (physical environment, resources). Teachers' negotiations of these components are influenced by the *iteration* dimension (teacher's personal and professional experience) and the *projective* dimension (their short- and long-term aims). Our adapted model considers the *practical-evaluative* and *projective* dimensions and CMs as a key material component. Our analysis focused on how CMs guide teaching decisions and how teachers described using this guidance to inform three levels of curricular aims: Macro-level refers to broad instructional goals and objectives; mid-level refers to sequencing and depth of topics included; lesson-level refers to the content taught in a lesson and the teaching techniques used. We drew on existing research to identify cultural and structural components relevant to the use of curriculum materials in each context.

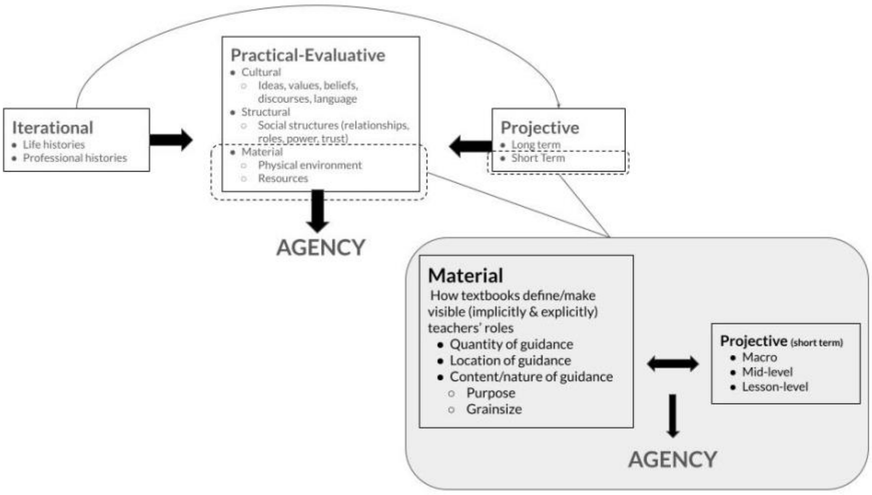


Figure 1. Ecological model of teacher agency as the interplay between CMs and teacher decision-making in relation to their instructional aims

CROSS-CULTURAL RESEARCH ON TEACHERS' USE OF CMS

Our examination of how elementary teachers in four cultural contexts describe their use of mathematics CMs brings together existing cross-cultural research on teaching practices in different education systems, comparative textbook analyses, and research on teachers' use of mathematics textbooks. Culturally specific patterns and approaches in mathematics teaching have been documented by a number of studies (e.g., Pepin & Haggarty, 2001; Stigler & Hiebert, 1999), although the research on elementary mathematics is limited. These differences have been explained by cultural norms and values (Stigler & Hiebert, 1999), which are reflected in the structures and policies of the particular education system (Krzywacki et al., 2023).

CM comparisons have also found cross-cultural differences, but the majority focus on the treatment of mathematics topics and offer limited insight into pedagogical practices or how textbooks are designed to guide teachers. One exception exists: An analyses of how mathematics CMs from Flanders, Sweden, and the U.S. communicate with teachers found differences in the mode and focus of communication and the positioning of teachers in relation to authority (Van Steenbrugge & Remillard, 2023).

A few studies have made strides toward connecting cultural differences reflected in textbooks and classroom practice. Pepin and Haggerty (2001) detailed how cultural values in England, France, Germany, and Norway are reflected in policy documents, mathematics textbooks, and teaching practices. Further, in an examination of Swedish and Finnish teachers' interactions with Finnish CMs, Koljonen (2020) found differences in lesson planning and enactment, reflecting distinct cultural norms in each country. The impact of CM designs and their usage on these cultural differences remains less explored. In many settings, textbooks are viewed as potential messengers of change (Hemmi et al., 2017). At the same time, it is well understood that teachers use CMs in different ways (Remillard, 2005). This study looks more closely at how a large sample of teachers describe using the distinct features of their mathematics CMs.

CULTURAL AND STRUCTURAL COMPONENTS

As Nordic countries, Finland and Sweden share many cultural and educational values, including teacher autonomy and local decision-making, which are reflected in educational norms and policies in both, although with some variations. In Finland, the educational system promotes student and teacher independence and holds teachers in high regard (Krzywacki et al., 2023). The Finnish national curriculum sets overall topics to be covered, and the decentralized system allows teachers to select materials and determine when and how these topics are taught.

Sweden also grants teachers liberty to choose and modify CMs. The Swedish national curriculum outlines goals and guidelines for mathematics content and pedagogical approaches, but schools determine daily instructional structures. Individualized learning is highly valued in Sweden; elementary teachers may use several CMs to cater to students at different levels. In both countries, CMs are commercially published, with substantial input from experienced teachers (Hemmi et al., 2017).

The Flanders region of Belgium has a semi-autonomous structure. The government adopts attainment targets that identify the knowledge, skills, and attitudes that students should reach at the end of grade six, but authority to determine how to reach them is given to schools. School principals, often in consultation with teachers, select CMs, which are developed by a mix of stakeholders, are commercially published, and reflect the attainment targets. Teacher professionalism, where teachers contribute meaningfully to educational aims, tends to be valued over autonomy (Simons & Kelchtermans, 2008). As such, teachers are expected to regularly use selected CMs.

Education in the U.S. places considerable emphasis on control and accountability in order to reach common outcomes (Krzywacki et al., 2023). Each state establishes specific learning objectives for each grade level, and students' attainment of these objectives is measured annually on high-stakes assessments (Remillard & Reinke, 2017). Commercial publishers typically write CMs, although some are prepared by researchers and sold commercially. District administrators are responsible for selecting CMs that all teachers in the district are expected to use, often following a set schedule.

METHODS

Survey development

Data for the study came from a mixed-methods study of elementary teachers' use of mathematics CMs in Finland, Flanders, Sweden, and the U.S., which included semi-structured teacher interviews and curriculum analysis. The survey was designed to situate interview findings in a broader, cultural context. Survey questions asked about print and digital materials, frequency and purpose of use, and influencing factors. The team generated questions in English (common language) based on emerging themes from interviews in all four contexts. Following standard survey design procedures, questions were translated into the native language of each context (Dutch, Swedish, and Finnish), field tested using cognitive interviews, and then revised by the full team.

Survey Administration

The survey was launched in Spring 2022 and was open for approximately 6 months. Efforts to solicit respondents varied by context, based on typical practices and the team's access to resources in each location. The final data set included 397 teachers, grades 1-6, as follows: Finland (n=27), Flander (n=86), Sweden (n=102), US (n=182).

Analysis

Responses were compiled, cleaned, and arrayed to facilitate comparative analysis. A survey consultant performed analysis of variance on appropriate responses to determine statistical significance of differences found. We identified questions relevant to the focus of this report–understanding patterns in how teachers use CMs to design lessons–and undertook descriptive analyses, looking for relevant patterns within or across contexts. To consider the role that CMs might play in these patterns, we identified the five most commonly named CMs in each context and analyzed them to describe patterns in the guidance CMs provide.

FINDINGS

Following our ecological model of teacher agency, we first present findings from the analysis of CMs in each context and then use survey findings to explore patterns in teachers' reported actions and decisions in relation to their CMs and their aims.

How CMs Structure Teachers' Roles

Analysis of the five most commonly mentioned CMs from each context surfaced relevant patterns that resonate with the cultural and structural components described earlier. For example, all CMs were aligned with national or state official curriculum documents, however, they differed in how specifically they mapped this overlap. Finnish and Swedish CMs offered macro-level frameworks over prescriptive guidance, prioritizing flexibility and autonomy in how teachers enact their roles. In contrast, the U.S. and Flanders CMs provided more mid- and lesson-level guidance, further specifying the teacher's role. We summarize these findings in the following paragraphs. We then detail survey findings to explore how teachers in each context interpret and act in relation to these features.

Finnish CMs analyzed were predominantly similar in structure and content, with a moderate level of guidance that aligns with the national core curriculum standards. predominantly similar in structure and content, providing a macro-level framework that encouraged teachers to integrate their own pedagogical approaches within the established educational goals. Each unit began with a description of the content and learning objectives. The daily lessons included student activities, but only moderate guidance for the teacher. This structure provided teachers freedom to decide on daily teaching practices, reflecting the core value of teacher autonomy.

The commonly used CMs identified by Swedish participants varied considerably and included both traditional and newly developed CMs. However, all CMs linked core content and objectives to mathematical concepts and skills developed through each unit or lesson; they also offered differentiated tasks to cater to individual needs. Like the Finnish CMs, and in keeping the value of teacher autonomy, the traditional Swedish CMs included limited guidance for teaching daily lessons. In contrast, newer CMs provide greater pedagogical guidance, suggesting a shift toward a more structured approach. Notably, the fourth most common response to the question, "What is your primary mathematics curriculum program?" was "I don't use a specific CM".

Flanders' CMs exhibit a structured and didactic approach, closely adhering to government-set attainment targets. Most CMs were detailed and prescriptive, providing teachers with step-by-step guidance for each lesson, including strategies for differentiation and adherence to structured learning plans. This high level of detail aims to ensure that instructional practices are consistent with the specified learning outcomes of the Flemish educational system.

Like Sweden, the U.S. CMs varied considerably. They were similar, however, in their comprehensiveness and explicit alignment with state standards, reflecting the U.S. emphasis on control and accountability. Each CM included extensive detail about sequence, scope, and learning objectives, providing considerable detail on how the components of each lesson should flow and specific pedagogical actions. Guidance also included suggestions for differentiation, prompts for discussion, and alerts about possible student misconceptions and how to address them.

Teachers' Decisions and the Interplay of CMs and Teachers' Own Ideas

Survey question asked teachers to report on factors that influenced their mathematics teaching. Responses from all contexts revealed that teachers characterized their curricular decisions as being influenced by their CMs and their own ideas. However, the degree to which teachers prioritized CMs or their own ideas differed across the contexts, highlighting differences in how teachers perceived their roles.

Teachers in all four contexts reported using CMs regularly ("almost every or most lessons") to a similar extent, ranging from 58% to 69%. Teachers from the U.S. and Finland, however, were more likely to indicate that their CMs were "very influential" in their mathematics teaching (63 and 62% respectively) than those from Flanders (36%) and Sweden (49%). This difference did not appear to be related to having adequate materials, since over 75% or more in each sample reported having adequate curriculum materials available for instruction.

Teachers in all contexts also reported relying substantially on their own ideas. A higher proportion of teachers from Nordic countries reported this position. Finnish teachers stood out, with 81% reporting that their own ideas were "very influential," compared to 72% of Swedish teachers, 52% of U.S. teachers, and 49% of Flanders teachers.

These findings surface several themes related to the interplay between objective components of professional space and teachers' enactment of agency: U.S. teachers stood out in their reliance on CMs when making curricular decisions. They were the only group to report being influenced by CMs at a higher rate than their own ideas. This pattern comports with the themes of structured guidance, accountability, and control present in the U.S. school environment. In contrast, teachers from Nordic countries prioritize their own ideas in their decisions. Finnish teachers stood out in the extent to which they report relying on their own ideas *and* CMs. Flemish teachers' responses suggest more discomfort in relying on CMs.

Teachers' Decisions in the Interplay of CMs and Curricular Aims

In light of the differences in how CMs in each context structured teachers' roles, we used survey data to explore *how* and *why* teachers from each context reported using CMs, and particularly the teacher's guide. Survey questions asked teachers to indicate how much they used their CMs to determine overall goals and objectives (macro-level), sequencing topics and time spent on each (mid-level), content and skills taught in each lesson and teaching techniques to be used (lesson-level). Our aim was to consider how teachers made use of objective features in their CMs to enact curricular aims, as well as how their approaches related to cultural values and structures in their systems.

It appears that the balance between the influence of CMs and teachers' own ideas, outlined above, reflects the structure and content of guidance in their CMs, along with cultural norms and values. U.S. CMs provided extensive and comprehensive guidance, and U.S. teachers were likely to report using them to determine all levels of curricular aims “primarily or quite a bit" (macro 74%, mid 72%, lesson content 70%, teaching techniques 59%). U.S. teachers also stood out in the extent to which they reported following CMs; 62% indicated they “follow[ed] the script in the teacher’s guide as written or mostly as written,” compared to 8-11% of teachers from European contexts.

Teachers in our study from Nordic countries reported more limited use of CMs and this use was concentrated around macro-level aims; just over half reported strong reliance on CMs for some macro-level purposes (61% Finland, 54% Sweden). Far fewer respondents reported relying on CMs for mid-level (32, 33%). At the lesson-level, these teachers were more likely to report using CMs to determine the content to be taught (53, 60%) than teaching techniques and activities (37, 36%). These findings reflect the nature of guidance provided by CMs and the cultural expectation that teachers make teaching determinations, using multiple CMs and resources.

Teachers in Flanders present a mixed profile that does not align cleanly with either pattern above. They were more likely than the others to report using CMs to determine macro-level aims (87%), and they fell in the middle on using CMs to determine mid-level aims (50%). Their reported use of CMs for lesson-level aims was akin to the U.S. teachers: 76% for content to be taught and 42% for teaching techniques and activities. It appears that, despite the prescriptive approach found in Flemish CMs, teachers make selective use of the guidance for their mathematics instruction.

DISCUSSION & CONCLUSION

Our analysis of teachers’ reported use of CMs in four contexts from an ecological perspective of teacher agency contributes to our understanding of teachers’ professional actions in relation to CMs within a cultural and institutional context. It is well understood that CMs are used around the world and operate differently in different school systems (Valverde et al., 2002). By considering teachers’ decisions and actions, our findings add complexity to what it means to “use” CMs. We found that teachers in different contexts negotiated the guidance in their materials, relying on them for different instructional purposes and in relation to their own ideas.

By adopting a professional space lens (Oolbekkink-Marchand et al., 2017), our analysis demonstrates how CMs operate as part of the objective components of school systems, reflecting and reinforcing cultural values and institutional structures. They are not simply messengers about mathematics curriculum, but about traditions and values related to teaching and teachers' roles.

Our findings offer some insight into how teachers interpret and negotiate these structures as actors in this system. Teachers made selective and purposeful use of components of CMs that were pertinent to their contexts and instructional aims. Using CMs to guide or inform instructional decisions was not understood as a rejection of one's own ideas and values. The Finnish data illustrates this perspective well, showing that teachers embraced being influenced by both CMs and their own ideas.

References

Hemmi, K., Krzywacki, H., & Koljonen, T. (2018). Investigating Finnish Teacher Guides as a Resource for Mathematics Teaching. *Scandinavian Journal of Educational Research, 62*(6), 911–928.

Krzywacki, H., Condon, L., Remillard, J. T., Machalow, R., Koljonen, T., & Steenbrugge, H. V. (2023). Emergency remote teaching as a window into elementary teachers' mathematics instructional systems in Finland and the U.S. *International Journal of Educational Research Open*, *5*, 100286.

Oolbekkink-Marchand, H. W., Hadar, L. L., Smith, K., Helleve, I., & Ulvik, M. (2017). Teachers' perceived professional space and their agency. *Teaching and Teacher Education*, *62*, 37–46.

Pepin, B., & Haggarty, L. (2001). Mathematics textbooks and their use in English, French and German classrooms: A way to understand teaching and learning cultures. *ZDM*, *33*(5), 158–175.

Priestley, M., Edwards, R., Priestley, A., & Miller, K. (2012). Teacher Agency in Curriculum Making: Agents of Change and Spaces for Manoeuvre. *Curriculum Inquiry*, *42*(2), 191–214.

Remillard, J. T. (2005). Examining key concepts in research on teachers' use of mathematics curricula. *Review of Educational Research*, *75*(2), 211–246.

Simons, M. & Kelchtermans, G. (2008). Teacher professionalism in Flemish policy on teacher education: a critical analysis of the Decree on teacher education (2006) in Flanders, Belgium, *Teachers and Teaching: theory and practice*, *14*(4), 283-294.

Stigler, J., & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom*. The Free Press.

Valverde, G. A., Bianchi, L. J., Wolfe, R. G., Schmidt, W. H., & Houang, R. T. (2002). *According to the book: Using TIMSS to investigate the translation of policy into practice through the world of textbooks*. Kluwer.

Van Steenbrugge, H., & Remillard, J. T. (2023). The multimodality of lesson guides and the communication of social relations. *ZDM – Mathematics Education*, *55*(3), 579–595.

Wermke, W., & Höstfält, G. (2014). Contextualizing Teacher Autonomy in time and space: A model for comparing various forms of governing the teaching profession. *Journal of Curriculum Studies*, *46*(1), 58–80.

[1]Note: We use the term curriculum materials (CMs) broadly to refer to teacher's guides, student textbooks, and other curriculum resources that are teacher facing.

TEACHERS' MOTIVATIONS TO TRANSITION TO DE-STREAMED SECONDARY MATHEMATICS

Kaitlin Riegel, David Pomeroy, Sara Tolbert, and Kay-Lee Jones

University of Canterbury

An initiative supported by the Ministry of Education to combat educational inequity has positioned de-streaming New Zealand secondary mathematics as a critical issue. Using the lens of self-determination theory (SDT), understanding teachers' beliefs, motivations, and goals in de-streaming may facilitate this transition. This report presents the results of a thematic analysis on interviews from 11 secondary mathematics teachers. Findings suggest that teachers can internalise goals of externally introduced de-streaming initiatives and teacher intrinsic motivation can originate from knowledge of the broader negative consequences of streaming. Practical implications are discussed together with the results.

INTRODUCTION

'Streaming', also known as 'setting' or 'tracking,' is the practice of grouping students by perceived ability, commonly through sorting them into separate classes (Domina et al., 2019). The removal of streaming, or 'de-streaming,' is a pressing issue in Aotearoa New Zealand due to evidence of inequitable outcomes, particularly for Māori and Pasifika students (Tokona Te Raki, 2019). These findings, together with cumulative evidence from international literature on the harms of streaming to equity (Terrin & Triventi, 2023) informed an action plan for de-streaming schools by 2030 with support of the Ministry of Education (Tokona Te Raki, 2023). However, de-streaming can be difficult and complex (Horn, 2006; Taylor et al., 2017), with a small but increasing portion of schools in the country currently de-streamed or transitioning to de-streamed classrooms (OECD, 2023). Mathematics particularly remains heavily streamed, which may contribute to postsecondary inequity, given its role as a 'gatekeeper' subject.

One explanation for the difficulty of effectively de-streaming, is that the motivations and goals of teachers may conflict with those of school leaders and researchers (Taylor et al., 2019). Thus, it is helpful to examine the motivation of teachers participating in a project aimed at understanding and supporting the transition to de-streamed secondary mathematics. Further, understanding teachers' goals for de-streaming with respect to their beliefs and motivation may inform supporting teachers in constructing and reaching appropriate goals. Doing so will enable those who are leading the transition in their own schools to support their teachers through this process.

LITERATURE AND THEORETICAL BACKGROUND

Beliefs heavily shape what is valued, which, in turn, influences motivation and goals (Parks & Guay, 2009). Many teachers believe that streaming has positive impacts on

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. 4, pp. 65–72). PME.

students (Taylor et al., 2017), particularly through being effective for students of varying prior attainment. This belief can be stronger in mathematics (Hallam & Ireson, 2003), particularly in higher year levels (Forgasz, 2010). Teachers view the disadvantages of streaming as including problems with stream allocation, and differing educational quality and opportunities between streams, both of which restrict long-term options (Forgasz, 2010). While teachers sometimes indicate the limitations of streaming for certain students, they often do not agree upon or acknowledge broader equity issues perpetuated by streaming (Forgasz, 2010; Hallam & Ireson, 2003). Importantly, school ethos around mixed grouping influences teacher beliefs (Hallam & Ireson, 2003). There is little research on the transition to de-streamed classrooms, leaving teachers beliefs, motivations, and goals through this shift poorly understood.

Self-determination theory (SDT) diverges from many previous theories of motivation, which focus on how external factors influence behaviour, through further centralising internal factors that shape development and regulation (Ryan & Deci, 2019). SDT distinguishes the ‘quality’ of motivation rather than simply motivational intensity (Deci & Ryan, 2008), which can correspond with wellbeing, performance, and persistence (Ryan & Deci, 2000). It is therefore an appropriate framework for this study since the transition to de-streaming requires a long-term commitment.

SDT posits that three psychological needs must be fulfilled to enhance self-motivation: *competence* (perceived effectiveness in relevant tasks), *autonomy* (personal agency), and *relatedness* (connectedness and belonging with others) (Deci & Ryan, 2008). *Intrinsic motivation*, participating in an activity for its inherent interest and satisfaction, is supported through meeting these needs (Ryan & Deci, 2019). *Extrinsic motivation* exists on a spectrum, which varies in the level of individual autonomy. At one end, there is *external regulation* and *introjected regulation*, where regulation takes place through purely external factors, such as compliance, or is not fully internalised. In contrast, *identified regulation* and *integrated regulation* can be understood as more closely linked to intrinsic motivation through individuals valuing or assimilating to the goal content (Ryan & Deci, 2000).

Goals refer to ‘what’ someone expects to accomplish from their behaviours, whereas motivations are ‘why’ they engage in the behaviour (Deci & Ryan, 2000). Goal content may be intrinsic or extrinsic, with intrinsic goals associated with autonomous motivation, wellbeing, and the satisfaction of the basic psychological needs (Deci & Ryan, 2000). Higher-quality motivation can correspond with stronger outcomes relevant to professional development. It is therefore important to unpack sources of motivation and identify related goals, to understand and support teachers in their transition to de-streamed mathematics. With this in mind, our research questions were:

- What motivates teachers to participate in a project aimed at supporting and understanding the de-streaming of secondary mathematics?
- How are teachers’ beliefs about streaming, motivations, and goals for participating in this project related?

METHODS

This research sits within a broader project that aims to understand and support the de-streaming transition in partnering with teachers, students, and schools. Five New Zealand schools, which had recently de-streamed secondary mathematics and represented regional, ethnic, and socioeconomic diversity, were recruited as partners. Eleven teachers from these schools participated in semi-structured interviews in-person or online. After listening to the interviews while making preliminary notes, they were coded by the first author using inductive thematic analysis (Braun & Clarke, 2012). The themes were reviewed separately by the second and third author, then refined. The results outline whether and how participants expressed their beliefs about streaming, their motivations for participating in this de-streaming project, and their goals in participating.

RESULTS

Beliefs

Of the participants who articulated beliefs about de-streaming, a clear subset was identified, who expressed the view that streaming limits educational pathways.

> Claire: Why do we get to put them in Year 11 into this class…? You're stopping them from getting UE (*university entrance*), you're stopping them from doing anything that has maths and statistics in their future, what right do we have to make that decision?

This was often linked to the algebra and calculus content of streams.

> Tom: I've talked for lots of years about how streaming isn't fair…limiting people and only teaching the extension class the higher-level stuff. What about all of these kids in the next sort of two or three bands, however you stream? They're missing out on that high-level stuff, and I know, I've seen it so many times that you struggle and struggle and struggle with algebra, and then 'click,' all of a sudden, it happens. And that happens at different times for everybody. So, why would we stop them from trying?

Teachers highlighted their beliefs about bias in streamed systems based on their own experiences. Explicitly linking to broader equity issues, the belief around systematic bias was particularly evident with respect to racial inequity.

> Tom: But our extension classes are full of the kids whose parents have fought for the right for their kids to be there…And that's going to link into our culturally responsive[ness] isn't it? You know, what do they say systemic racism, you know, yeah that's the other side.

Diane elaborated further on her perspective of de-streaming facilitating positive outcomes for students stating, "Our achievement in external [exams] is going up, but the bar was quite low. Our engagement's gone up because we said everybody's doing externals now."

Motivations

Contrasting motivations for participating in the project distinguished the teachers. One extrinsic motivation was being recruited for participation by school leadership or colleagues. This recruitment existed on a spectrum and was sometimes a clear top-down directive,

> Claire: [The head of mathematics] said we're changing everything. We're going to do this and this, and being a second-year teacher I was like 'okay,' and just went along with everything…now it's something that I am really passionate about, but not because of [colleague] but because of me.

For some, their engagement was presented as optional,

> Ivan: My head of department asked me if I wanted to join…and I thought, yeah that could be something I'm into…so just pretty much said yes immediately and it wasn't until I suppose we had our first meeting…I realised that, oh, this is actually something I really agree with.

Some were motivated by their peers,

> Marama: I find [colleague] very inspirational. Just her drive and passion and for the kids to be successful and to love what they do when they're in her classroom…Yeah, I get caught up in that, I guess. I look at the impact on our kids and that's inspiring for me.

Similarly, some teachers commented on the project aligning with their schools' goals. Extrinsic motivation can take the form of a directive or feeling pressure to conform. However, some extrinsically motivated teachers began to value or internalise the goals of de-streaming. Several teachers identified such extrinsic sources as their only motivation, while others reported other factors. Many participants were motivated by their observations of the inequity, injustice, and damage of streaming.

> Kyle: I think the real goal behind that is creating an inclusive and relational culture where there's diversity and students will feel included and students feel empowered to be part of, where students feel actually part of the kura [school] they feel like they belong.

Such concerns were linked to personal reflections on teachers' own education.

> Brie: I was in extension classes. I was in the top groups…But when I think back now it's actually because the system worked for me, not because I was intelligent…I think teaching changed that view a lot because you see so much more. You see all the different kids coming in from all the different places then there's not much equity there.

Student and societal attitudes toward mathematics and education generally were raised as a motivator for engaging with this project. This, again, was often linked to reflections on their own educational journey.

> Diane: I feel very passionate about de-streaming from my own personal journey of being streamed as a child myself at high school with English and I still have

> whakamā [shame/embarrassment] around English…I feel like if that's the narrative that we're perpetuating for these students from day one, it's just so hard to change any of their dispositions with themselves in mathematics, let alone in themselves as a learner in our system.

Some teachers were intrinsically motivated by their curiosity in or challenge of de-streaming mathematics at secondary level.

Goals

Goals for participation were broad, though nearly all teachers indicated a desire for clarity, guidance, or resources for facilitating or improving their de-streamed practices. Teachers sought to gain a better understanding of the effects of streaming and what it means to de-stream effectively.

> Diane: What I professionally hope we can get out of it is looking more at the pedagogy and kaupapa [purpose] that needs to be happening in the classroom…So, we've got the de-streaming part sorted, but now it's like actually how do we support this happening in the classrooms? How do we support our kaiako [teachers]? How do we support our ākonga [students] to make sure that they're still getting the best out of what we can offer?

Teachers hoped the project would provide the opportunity to expand their toolbelt with concrete approaches to inform their practice. Tania comments, "I'm just looking at what are the different tools, or what are the different strategies we can use. And I think we need to create our own, you know, unique style of going about doing it."

Teachers sought guidance to inform their own practice, including outside the classroom. Layla commented on managing parental responses to de-streaming and Ivan sought to learn how to convince other staff of its value. No teachers discussed these goals outside of intrinsic personal and professional growth to support their students. For example, there was no reference to their practices needing to reflect well on their classrooms. In participating, many hoped to not only gain knowledge from others, but also form collaborations. Tom explained, "I've got some ideas and I like to bounce those ideas off others…being able to talk to others and hear how other people are approaching similar concerns, similar issues."

However, some teachers expressed goals that exceeded the boundaries of developing their own understanding and transition to de-streamed classrooms. Specifically, teachers hoped to support equity and cultural empowerment. Marama elaborated, "I guess what I would like is for our rangatahi Māori [Māori youth] to feel confident to take pāngarau [mathematics] and science and technology and those subjects when they get into the senior school." This hope was linked to teachers' own classroom practices.

> Brie: I've realised that there's a lot of things we can change in our practice, not just to make things more equitable for kids that have other stuff going on, but actually just to make things more equitable for cultural differences and just how kids work… particularly around culturally empowering practices for our Māori ākonga [students].

Some hoped their participation would facilitate more widespread awareness and adoption of de-streaming. Kyle explains, “I’d really like to see more schools de-stream and…more teachers getting involved in thinking about, rethinking the way we do things a little bit.” This goal was linked to equity concerns around streaming, with Diane stating, “I think that the more widely we can get our ideas and thoughts and kōrero [conversation] and use data to show people how harmful the streaming can be, the better for everybody.”

Relationships between Beliefs, Motivations, and Goals

Three teachers did not express particular beliefs about streaming in their interviews and exclusively reported being motivated by leadership or other faculty members. These teachers discussed their goals for this project as seeking clarity around de-streaming, exemplars of successful practices, resources for their own work, and collaborators in their journey. Two other teachers, who expressed the negative beliefs they held about streamed systems and discussed only intrinsic sources of motivation regarding equity and attitudes toward mathematics, extended on these goals to include improving equitable practice and supporting widespread adoption of de-streaming.

Four teachers reported both extrinsic and intrinsic sources of motivation for their participation. Three of these identified goals that overlapped with the intrinsically motivated teachers. However, despite acknowledging inequity as a belief and motivation, for the fourth teacher, addressing societal change was not identified as a goal. One teacher was intrinsically motivated by the challenge of de-streaming and also acknowledged inequity as a motivating factor, but did not discuss beliefs about streaming. Her goals were similar to those of the extrinsically motivated teachers. Two Māori teachers, who worked in schools with a high proportion of Māori students, did not explicitly express beliefs on streaming. Their semi-structured interviews largely focused on the cultural empowerment of their students, with de-streamed practice treated by them as a prerequisite connected to their participation in the project.

DISCUSSION

Many schools across New Zealand will be moving to de-streamed mathematics, so for teachers the transition will be, to some extent, extrinsically motivated. To facilitate an effective transition, teachers will ideally develop identified or integrated regulation (Ryan & Deci, 2000). Teachers in this project participated voluntarily, but there were still varying sources of motivation for their engagement. Extrinsic sources came on a spectrum from being introduced by leadership, to being exposed by inspiring colleagues, as well as school values that had been personally adopted. Intrinsic sources of motivation were identified to be inequity, students’ attitude toward mathematics, and inherent interest in shifting practice. Our analysis suggests that it is realistic for teachers to internalise the goals of externally introduced de-streaming initiatives. For schools moving to de-stream, a source of intrinsic motivation appears to often originate from teachers’ concerns about the broader societal inequity and negative attitudes toward mathematics, which streaming can perpetuate. The influence of broader societal inequity differs from previous findings (Forgasz, 2010) and exposure to this information may be a critical component of the transition.

The teachers' goals for participating in this project ranged from practical insights, such as guidance and resources, to broader issues, such as equity, cultural empowerment, and encouraging de-streaming nationally. The content of these goals can be understood as intrinsic in their professional development and personal satisfaction, through encouraging meaningful change for others. It will be important to allow teachers to develop appropriate intrinsic goals during the transition. Therefore, teachers should not be forced into formulating extrinsic goals, which can undermine motivation (Ryan & Deci, 2000), for example, through tracking student achievement while transitioning.

In line with SDT, the teachers' motivations and goals in this study associated with fostering their basic psychological needs. Goal content including guidance for improving their practices connects to their competence and sense of autonomy within their de-streamed classrooms. Developing goal content around fostering a community of collaborators corresponds with relatedness. Our analysis suggests that for all teachers in their de-streaming journey, these are important goals for schools to help address to encourage higher-quality motivation. However, some teachers had broader motivations and goals in their participation, specifically, going beyond the satisfaction of their basic psychological needs and pertaining to the development of equity and cultural empowerment nationally. The desire and intention to do good for others may be a powerful need to be satisfied in this context to drive goal content and motivation, in addition to the basic psychological needs posited by SDT (Martela & Riekki, 2018).

Our findings suggest that more autonomous motivation is evident when valuing and developing goals around educational equity are apparent. Thus, encouraging a school-wide dialogue around broader social issues and presenting this as a goal of de-streaming is a potential approach in fostering intrinsic teacher motivation for de-streaming, though this may not be completely internalised for everyone early in the process. The notion that de-streaming may be considered as a prerequisite for the goal of developing culturally empowering practice will be explored in future analyses.

References

Braun, V., & Clarke, V. (2012). Thematic analysis. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *APA handbook of research methods in psychology, Vol. 2. Research designs: Quantitative, qualitative, neuropsychological, and biological* (pp. 57–71). American Psychological Association. https://doi.org/10.1037/13620-004

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*(4), 227-268. https://doi.org/10.1207/S15327965PLI1104_01

Deci, E. L., & Ryan, R. M. (2008). Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian Psychology, 49*(3), 182-185. https://doi.org/10.1037/a0012801

Domina, T., McEachin, A., Hanselman, P., Agarwal, P., Hwang, N., & Lewis, R. W. (2019). Beyond tracking and detracking: The dimensions of organizational differentiation in schools. *Sociology of Education, 92*(3), 293-322. https://doi.org/10.1177/0038040719851879

Forgasz, H. (2010). Streaming for mathematics in years 7-10 in Victoria: An issue of equity? *Mathematics Education Research Journal, 22*(1), 57-90. https://doi.org/10.1007/BF03217559

Hallam, S., & Ireson, J. (2003). Secondary school teachers' attitudes towards and beliefs about ability grouping. *British Journal of Educational Psychology, 73*, 343-356. https://doi.org/10.1348/000709903322275876

Horn, I. S. (2006). Lessons learned from detracked mathematics departments. *Theory Into Practice, 45*(1), 72-81. https://www.jstor.org/stable/3497019

Martela, F. M., & Riekki, T. J. J. (2018). Autonomy, competence, relatedness, and beneficence: A multicultural comparison of the four pathways to meaningful work. *Frontiers in Psychology, 9*, 1157. https://doi.org/10.3389/fpsyg.2018.01157

OECD. (2023). *PISA 2022 Results (Volume II): Learning During – and From – Disruption*. PISA, OECD Publishing. https://doi.org/10.1787/a97db61c-en

Parks, L., & Guay, R. P. (2009). Personality, values, and motivation. *Personality and Individual Differences, 47*(7), 675-684. https://doi.org/10.1016/j.paid.2009.06.002

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68-78. https://doi.org/10.1037110003-066X.55.1.68

Ryan , R. M., & Deci, E. L. (2019). Brick by brick: The origins, development, and future of self-determination theory. In A. J. Elliot (Ed.), *Advances in motivation science* (pp. 111-156). Elsevier. https://doi.org/10.1016/bs.adms.2019.01.001

Taylor, B., Francis, B., Archer, L., Hodgen, J., Pepper, D., Tereshchenko, A., & Travers, M. (2017). Factors deterring schools from mixed attainment teaching practice. *Pedagogy, Culture & Society, 25*(3), 327-345. https://doi.org/10.1080/14681366.2016.1256908

Taylor, B., Francis, B., Craig, N., Archer, L., Hodgen, J., Mazenod, A., Tereshchenko, A., & Pepper, D. (2019). Why is it difficult for schools to establish equitable practices in allocating students to attainment 'sets'? *British Journal of Educational Studies, 67*(1), 5-24, https://doi.org/10.1080/00071005.2018.1424317

Terrin, É., & Triventi, M. (2023). The Effect of School Tracking on Student Achievement and Inequality: A Meta-Analysis. *Review of Educational Research, 93*(2), 236-274. https://doi.org/10.3102/00346543221100850

Tokona Te Raki. (2019). *He Awa Ara Rau - A Journey of many paths*. https://www.maorifutures.co.nz/research/research-item-old/

Tokona Te Raki. (2023). *Kōkirihia*. https://www.maorifutures.co.nz/projects/streaming/

GEOMETRY LEARNING OF STUDENTS WITH GENERAL LEARNING DIFFICULTIES: AN EYE-TRACKING STUDY ON THE IDENTIFICATION OF QUADRILATERALS

Maike Schindler[1], Anna Lisa Simon[1], Elisabeth Czimek[1], Benjamin Rott[1], and Achim J. Lilienthal[2]

[1]University of Cologne, Cologne, Germany, [2]TU Munich, Munich, Germany

Geometry is an important mathematical domain, especially for students with general learning difficulties (LD). However, not much is known about geometry learning of students with LD, possible difficulties, and needs for support. The aim of this paper is to investigate if and how students with LD differ in the identification of quadrilaterals from students without LD. We carried out an eye-tracking study with 184 students (20 with LD, 164 without LD) in which students were asked if given shapes were quadrilaterals. We analyzed students' error rates (from their oral responses) and their strategies, based on qualitative analysis of eye-tracking videos. We found that students with LD tended to make more mistakes than students without LD and to regard the quadrilaterals more often holistically, paying less attention to their properties.

INTRODUCTION

Geometry is one of the central domains of mathematics learning on primary and secondary school level (KMK, 2022). It is particularly important for students with general learning difficulties (LD): students who encounter learning difficulties that are severe, long-lasting, and extensive (e.g., Heimlich, 2016; OECD, 2007) affecting several school subjects, predominantly reading, writing, mathematics, and "learning to learn" (Heimlich, 2016, p. 36). For students with LD, geometry is crucial since, on the one hand, it is vital for everyday requirements and, on the other hand, it is central to relevant professional fields for students with LD (Basendowski & Werner, 2010; Hellmich, 2016). Due to the significance of geometry for the lives of students with LD, it is important for teachers to help them develop an understanding of geometric concepts and to support them adequately (Hellmich, 2016). However, there is a research gap on the geometry learning of students with LD and possible needs for support.

This paper addresses the geometry learning of students with LD. It focuses on one particular activity that is associated with supporting and investigating students' geometric understanding: the identification of geometric shapes, especially of quadrilaterals (e.g., Clements et al., 1999; Hannibal, 1999). In a study involving students with LD and students without LD, we pursued the research question, *Do students with LD differ from students without LD in the identification of quadrilaterals, and how?* We analyzed students' error rates (from their oral responses) and their strategies, based on qualitative analyses of eye-tracking videos.

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. 4, pp. 73–80). PME.

RELATED WORK

Learning difficulties and special educational needs in learning

The phenomena of learning difficulties and special educational needs are not clearly defined; different definitions and terms are being used, which are influenced by culture and educational systems, and there is often an interplay of several factors that lead to learning difficulties (Grünke & Canvendish, 2016). In short, in this paper – in line with our national educational system and its cultural context – we address *students with special educational needs in their learning* as students who encounter learning difficulties that are severe, long-lasting, and extensive (e.g., Heimlich, 2016; OECD, 2007). Students with special educational needs in their learning have difficulties in their academic achievement and can have difficulties in several school subjects, but usually show difficulties in reading, writing, mathematics, and in “learning to learn” (Heimlich, 2016, p. 36).

The few studies on the mathematical learning of students with LD indicate that students with LD show deficiencies in basic numerical competencies already at the beginning of school (Moser Opitz, 2008) and that these deficiencies appear to increase over the course of schooling: Werner et al. (2019) found that students with LD in grade 6 did not meet the level of mathematical competencies that are expected at the end of grade 4 (e.g., regarding arithmetic or geometry), while Gebhardt et al. (2015) found that the majority of fifteen-year-old students with LD were at or below the competence level I in PISA. Even by the end of school, students with LD often show difficulties with primary school level competencies, e.g., arithmetic operations, and often do not have the mathematical competencies required for the transition to working life (e.g., Lutz et al., 2023). While the existing studies indicate difficulties and delays (often with a product-view), to the best of our knowledge, there are no studies focusing specifically on the geometry learning of students with LD and possible needs for support.

Geometry and identification of shapes

The development of students’ understanding of geometric concepts is a complex process (Tsamir et al., 2008) and often described in simplified terms using van Hiele’s stage theory (van Hiele, 1986), which describes levels of geometric thinking “from a Gestalt-like visual level through increasingly sophisticated levels of description, analysis, abstraction, and proof” (Clements & Battista, 1992, p. 426, see also Clements et al., 1999). Van Hiele’s model includes five levels, of which the first three are relevant to this study: Level 0, “Visualization”, where students regard shapes through their holistic appearance, which is prototype-based; Level 1, “Analysis”, where students pay attention to properties of shapes and can use them to classify shapes; Level 2, “Abstraction”, where properties are being ordered and where relationships between properties can be used for reasoning. While younger children rely on prototypical ideas when identifying geometric shapes, property-based considerations gain significance as development progresses (Clements et al., 1999): While young learners (aged 4-6) identify prototypical quadrilaterals more easily than atypical ones and are influenced by mathematically irrelevant attributes such as shape orientation or aspect ratio (Unterhauser & Gasteiger, 2018), students from grades 1 to 4 increasingly rely on properties, although still prototypical ideas play a role for older students (Bruns et al., 2021). Fujita (2012) found that even among older students (aged 14), more than half of them relied on prototypes to identify quadrilaterals. All in all, quadrilaterals appear to be more difficult to identify than other shapes such as triangles (e.g., Ma, 2015).

Eye tracking

Eye tracking (ET) allows recording of spatio-temporal sequences of gaze points that indicate visual attention (Holmqvist et al., 2011). For mathematics education, ET is of interest, since the recorded sequences of gaze points allow inferences about cognitive (and affective) processes, although the interpretation of eye movements is not straightforward and bijective (e.g., Schindler & Lilienthal, 2019). ET has been of growing interest in recent years, partly because ET devices have become more affordable, advanced, accurate, and easy to use with students and in classroom settings (Lilienthal & Schindler, 2019; Strohmaier et al., 2020). ET offers benefits for mathematics education, especially for investigating processes of students with difficulties: In a methodological study comparing think-aloud interviews and ET analyses, Schindler and Lilienthal (2018) found that especially for students with mathematical difficulties ET offered more detailed insights, possibly because aspects such as “anxiety, difficulties with memory retrieval, introspection, or meta-cognitive reflection, or verbalization issues” (p. 117) may have influenced the verbal reporting of students with mathematical difficulties. Since such issues may also play a role for students with LD, our study uses ET to investigate their shape identification processes.

THIS STUDY

Participants

184 students who had just finished German primary school level and were in the first weeks of grade 5 participated in this study: 20 students with general LD who attended a special school for students with LD (mean age: 11.6 years) and 164 students attending an inclusive comprehensive school who did not have general LD (mean age: 10.7 years). The 20 students with LD were from three different school classes with different mathematics teachers each, while the 164 students without LD were from six classes with different mathematics teachers.

Setup

Data collection took place in individual sessions in a quiet room in the schools. The students were presented with a series of 32 shapes (prototypical quadrilaterals, non-prototypical quadrilaterals, and non-quadrilaterals) on a 24" full HD screen. The shapes were arranged in a random order that was the same for all students. Between tasks, children looked at a fixation star on the left side of the screen, which ensured a clear transition from one task to the next. The students were asked if the depicted shape was a quadrilateral and to answer “yes” or “no” as soon as they knew the answer. They received no feedback on their answers. The answers were recorded using an audio recording device. The eye movements were recorded with the Tobii Pro X3-120 eye tracker (120 Hz, binocular, infrared). This screen-based eye-tracker was attached to the bottom of the screen on which the shapes were presented. The average accuracy in our study was 0.9°. A 5-point calibration was performed with every student individually.

The presented shapes included the following three kinds (see Fig. 1, see also Tsamir et al., 2008, Simon et al., 2021): (1) *Prototypical quadrilaterals,* that means squares in prototypical orientation; (2) *Non-prototypical quadrilaterals*, that means atypical representatives, such as trapezoids or rhombuses; (3) *Non-quadrilaterals,* that means shapes that have a quadrilateral-like shape, yet, critical attributes of a quadrilateral are violated (e.g., curved or interrupted lines).

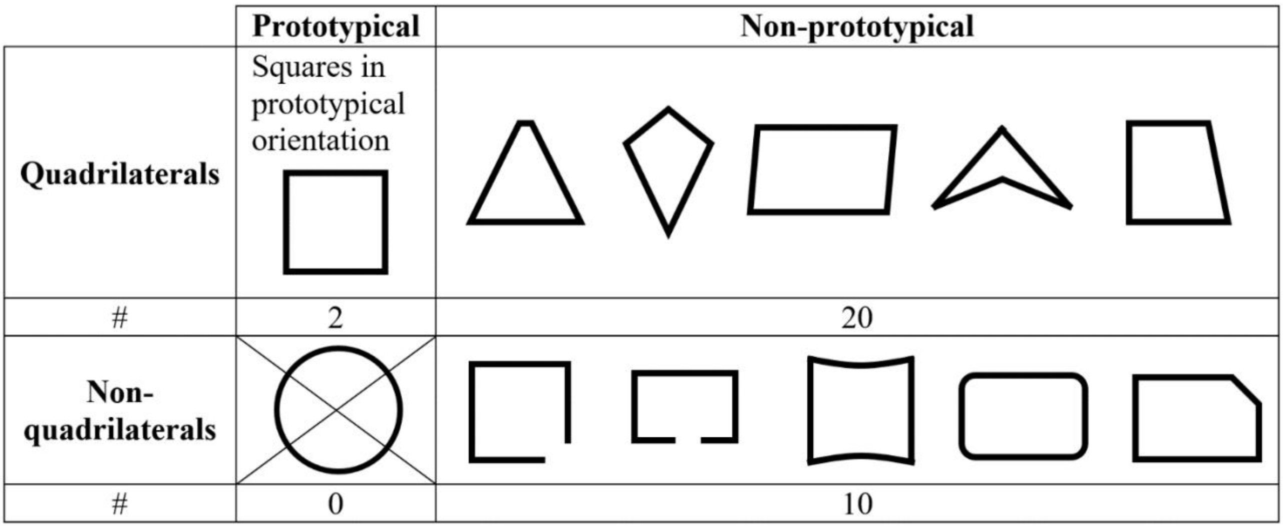


Figure 1. Item overview and examples (# indicates number of items)

Data analysis

Based on the students' responses, we determined the error rates. For investigating group differences regarding error rates, we carried out a repeated measurement ANOVA with type of measure (prototypical quadrilaterals, non-prototypical quadrilaterals, non-quadrilaterals) as within factor and group (with LD, without LD) as between factor. Due to the different number of items grouped under the different types of kinds of shapes, the relative error rates were used for the analyses.

We analyzed gaze-overlaid videos provided by Tobii Pro Lab software, where the gaze is visualized as a moving dot. For the qualitative analysis of the videos, the category system of strategies presented in Simon et al. (2021) was applied (see Fig. 2 and note that we use a static scanpath visualization (here: gaze plot) in this figure while the analysis was performed on ET videos). In short, the category system includes three kinds of strategies: (1) identifying *"at a glance"*, where the gaze moves to one aspect only (mostly the center of the shape or one vertex/side); (2) looking at *"parts of the shape"*, where gaze moves to distinct parts of the shape (e.g., two vertices, or a side and a vertex), yet, not the entire shape; (3) looking at the *"entire shape"*, where the gaze moves over the whole shape, partially multiple times (see Simon et al., 2021, for a more detailed description). To calculate interrater reliability using Cohen's kappa, 26% of the data were coded by a second coder independently. The interrater agreement was 0.87, which can be considered almost perfect (Landis & Koch, 1977).

For investigating group differences regarding students' shape identification strategies, we performed chi-squared tests based on the frequencies of strategy use for each kind of shape (prototypical quadrilaterals, non-prototypical quadrilaterals, non-quadrilaterals) and calculated effect sizes using Cramér's *V*. For prototypical quadrilaterals, we calculated Fisher-Freeman-Halton test, since the assumption of chi-squared tests was violated.

We used the software IBM SPSS 29 for the statistical analyses.

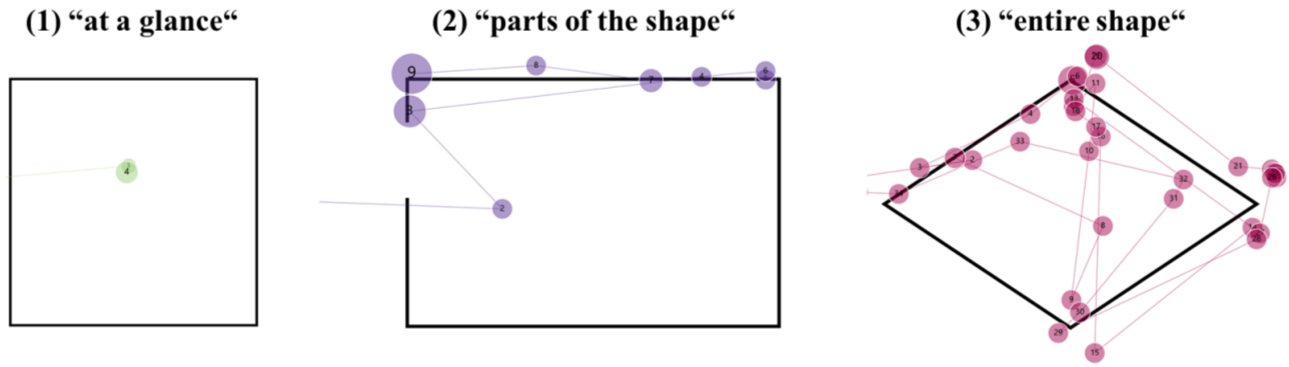


Figure 2. Kinds of strategies with scanpath examples

RESULTS

The analysis of *error rates* indicates that the students with LD in our study made more errors than students without LD: There was a significant main effect for the groups, independent of the types of quadrilaterals ($F(1, 182) = 6.08$, $p = .015$, $\eta p^2 = .03$). While *prototypical quadrilaterals* were recognized correctly as quadrilaterals by both groups of students with very few errors, *non-prototypical quadrilaterals* caused many errors for both groups (approx. 65% vs. approx. 60%). For *non-quadrilaterals*, students with LD made more errors than students without LD, i.e., they stated more often that the shapes were quadrilaterals (approx. 26% vs. approx. 15%): This means that for non-quadrilaterals, group differences in errors were most apparent.

We analyzed students' *strategy use* for prototypical quadrilaterals, non-prototypical quadrilaterals, and non-quadrilaterals separately (Fig. 3): For *prototypical quadrilaterals*, Fisher-Freeman-Halton test revealed a significant difference in strategy use between students with and without LD ($p = .047$, $V = .12$). For *non-prototypical quadrilaterals* and for *non- quadrilaterals*, chi-squared tests also revealed significant group differences in strategy use ($\chi^2\,(2) = 8.89$, $p = .012$, $V = .05$ and $\chi^2\,(2) = 6.20$, $p = .045$, $V = .06$). For all kinds of shapes, the students with LD tended to regard the shapes “at a glance” more often and tended to look at “parts of the shape” less often than students without LD (Fig. 3).

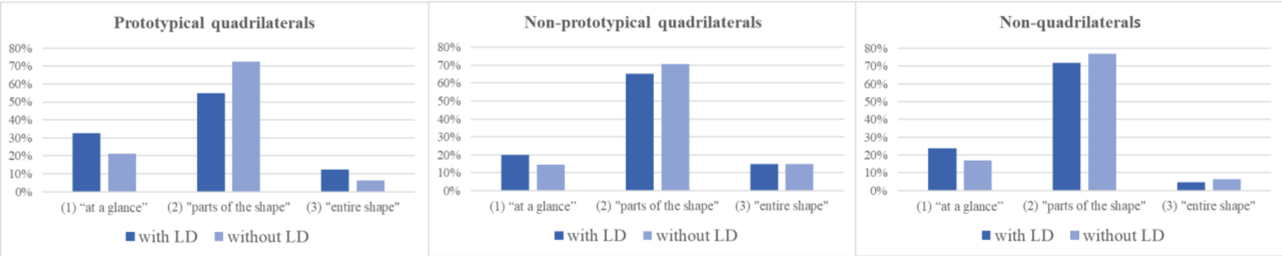


Figure 3. Strategy use for three kinds of quadrilaterals in this study

DISCUSSION

The aim of this paper was to investigate if and how students with LD differ in the identification of quadrilaterals from students without LD. We conducted an eye-tracking study and analyzed shape identification strategies from ET videos together with error rates in shape identification based on student oral responses.

When discussing the results, some limitations of the study need to be considered. First, the number of students with LD in our study was small, especially considering that students with LD are not a homogeneous group who cannot be assumed to be homogeneous in their mathematics learning either. The results of our study offer interesting insights, but future research should confirm them with larger groups and with students with LD from different school types, e.g., from special education schools (as in this study) and also from inclusive education. Second, it must be noted that although the group differences in our study were statistically significant, the effect sizes were small, indicating that the differences in strategy use and error rates between students with LD and without LD were practically small. Yet, our findings indicate meaningful trends, which we will summarize and discuss in the following.

In this study, we investigated strategies in identifying quadrilaterals. What we found is that students with LD for all three kinds of shapes (i.e., prototypical quadrilaterals, non-prototypical quadrilaterals, non-quadrilaterals) tended to use a quick strategy more often than students without LD: They tended to look at the shapes “at a glance” more often, which means that their gazes went to one point (often the center) and that they perceived the rest of the shape peripherally, while students without LD tended to look at parts of the shapes more often, paying more attention to the properties of the shapes (e.g., vertices, an interrupted or curved line). This goes along with the tendency of students with LD to make more errors in assessing whether a shape was a quadrilateral or not. Our findings indicate that the students with LD tended to perceive the shapes holistically more often and to pay (visual) attention to their properties less often than students without LD, which indicates that they more likely relied on prototypical ideas than students without LD. In particular, we saw that students with LD often regarded prototypical quadrilaterals at one glance and correctly, while this quick strategy went along with wrong answers for non-prototypical quadrilaterals. Students with LD also appeared to apply prototypical ideas more often for non-quadrilaterals, possibly overlooking or not regarding the violated critical attributes (e.g., when sides were curved), which in turn let them accept these shapes as quadrilaterals more often than students without LD. These findings connect to studies indicating that especially younger students rely on prototypical ideas but that they may persist even for older students (Unterhauser & Gasteiger, 2018; Bruns et al., 2021; Fujita, 2012) and to research that indicated differences in strategy use of students with and without LD (Grobecker & de Lisi, 2000). However, we found that also students without LD showed difficulties identifying non-prototypical quadrilaterals, which indicates that many students need support in this area. This is in line with research showing that quadrilaterals appear to be difficult for students to identify (e.g., Ma, 2015).

Finally, we saw in this study that ET provided detailed insights into the shape identification strategies of students with and without LD, without interfering with their strategies in a way think-aloud might have: Especially for students with LD, ET has great potential since the requirement to describe their strategies might have changed how the students with LD approached the task. Future research could build on this and use ET to inquire into students' strategy use also in longitudinal settings and to evaluate teaching or targeted support for students in the identification of geometric shapes.

References

Basendowski, S. & Werner, B. (2010). Die unbeantwortete Frage offizieller Statistiken: Was machen Förderschülerinnen und -schüler eigentlich nach der Schule? Ergebnisse einer regionalen Verbleibsstudie von Absolventen mit sonderpädagogischem Förderbedarf Lernen. *Empirische Sonderpädagogik, 2*, 64–88.

Clements, D. H. & Battista, M. T. (1992) Geometry and spatial reasoning. In Grouws, D.A. (Ed.), *Handbook of Research on Mathematics Teaching and Learning* (pp. 420–464). Macmillan.

Clements, D. H., Swaminathan, S., Zeitler Hannibal, M. A., & Sarama, J. (1999). Young children's concepts of shape. *Journal for Research in Mathematics Education, 30*(2), 192–212.

Fujita, T. (2012). Learners' level of understanding of the inclusion relations of quadrilaterals and prototype phenomenon. *The Journal of Mathematical Behavior*, *31*(1), 60–72.

Gebhardt, M., Sälzer, C., Mang, J., Müller, K., & Prenzel, M. (2015). Performance of students with special educational needs in Germany: Findings from programme for international student assessment 2012. *Journal of Cognitive Education and Psychology, 14*(3), 343–356.

Grobecker, B. & De Lisi, R. (2000). An investigation of spatial-geometrical understanding in students with learning disabilities. *Learning Disability Quarterly, 23*(1), 7–22.

Grünke, M., & Cavendish, W. (2016). Learning disabilities around the globe: Making sense of the heterogeneity of the different viewpoints. *Learning Disabilities: A Contemporary Journal*, *14*(1), 1–8.

Hannibal, M. A. (1999). Young children's developing understanding of geometric shapes. *Teaching Children Mathematics, 5*(6), 353–357.

Hellmich, F. (2016). Lehren und Lernen im Geometrieunterricht. In U. Heimlich & F. B. Wember (Eds.), *Didaktik des Unterrichts im Förderschwerpunkt Lernen: Ein Handbuch für Studium und Praxis* (pp. 294–306). Kohlhammer.

Heimlich, U. (2016). *Pädagogik bei Lernschwierigkeiten*. UTB.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. OUP.

KMK (2022). Bildungsstandards für das Fach Mathematik. Primarbereich. kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2022/2022_06_23-Bista-Primarbereich-Mathe.pdf

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174.

Lilienthal, A.J., & Schindler, M. (2019). Eye tracking research in mathematics education: A PME literature review. In M. Graven, H. Venkat, A.A. Essien, & P. Vale (Eds.), *Proceedings of 43rd Conference of the IGPME* (Vol. 4, p. 62). PME.

Lutz, S., Ebenbeck, N., & Gebhardt, M. (2023). Mathematical skills of students with special educational needs in the area of learning (SEN-L) in pre-vocational programs in Germany. *International Journal for Research in Vocational Education and Training, 10*(1), 1–23.

Ma, H. L., Lee, D. C., Lin, S. H., & Wu, D. B. (2015). A Study of Van Hiele of Geometric Thinking among 1st through 6th Graders. *Eurasia Journal of Mathematics, Science and Technology Education*, *11*(5), 1181–1196.

Moser Opitz, E. (2008*). Zählen, Zahlbegriff, Rechnen. Theoretische Grundlagen und eine empirische Untersuchung zum mathematischen Erstunterricht in Sonderklassen*. Haupt.

OECD (2007). Students with Disabilities, Learning Difficulties and Disadvantages. Policies, Statistics and Indicators. https://www.oecd-ilibrary.org/education/students-with-disabilities-learning-difficulties-and-disadvantages_9789264027619-en

Schindler, M., & Lilienthal, A.J. (2019). Domain-specific interpretation of eye-tracking data: Towards a refined use of the eye-mind hypothesis for the field of geometry. *Educational Studies in Mathematics, 101,* 123–139.

Schindler, M., & Lilienthal, A.J. (2018). Eye-tracking for studying mathematical difficulties—also in inclusive settings. In E. Bergqvist, M. Österholm, C. Granberg, & L. Sumpter (Eds.), *Proc. of 42nd Conference of the IGPME* (Vol. 4, pp. 115-122). PME.

Simon, A.L., Rott, B., & Schindler, M. (2021). Identification of geometric shapes: An eye-tracking study on triangles. In M. Inprasitha, N. Changsri, & N. Boonsena (Eds.), *Proc. of 44th Conference of the IGPME* (Vol. 4, pp. 47–55). PME

Strohmaier, A. R., MacKay, K. J., Obersteiner, A., & Reiss, K. M. (2020). Eye-tracking methodology in mathematics education research: A systematic literature review. *Educational Studies in Mathematics, 104,* 147–200.

Tsamir, P., Tirosh, D., & Levenson, E. (2008). Intuitive nonexamples: The case of triangles. *Educational Studies in Mathematics, 69*(2), 81–95.

Unterhauser, E., & Gasteiger, H. (2018). Verständnis des geometrischen Begriffs Viereck bei Kindern zwischen vier und sechs Jahren. *Frühe Bildung*, *7*(3), 152–158.

van Hiele, P. M. (1986). *Structure and insight: A theory of mathematics education*. Academic Press.

EFFECTS OF TEACHING STUDENTS TO SOLVE OPEN MODELLING PROBLEMS ON UTILITY, INTRINSIC, AND ATTAINMENT VALUES

Stanislaw Schukajlow[1], Janina Krawitz[2], Katharina Wiehe[1], and Katrin Rakoczy[3]

[1]University of Münster, Germany; [2]University of Paderborn, Germany; [3]University of Gießen, Germany

Task values are important for learning. However, prior research has indicated a lack of studies that have addressed students' task values in mathematics. In the following study (N = 293), we analyzed (1) the relationships between intrinsic, attainment, and utility values and (2) how teaching students to solve open modelling problems affects these values. Students in the experimental group were taught how to solve open modelling problems, whereas those in the control group were taught how to solve real-world problems with no missing information. Students reported their values before and after the intervention. The results revealed positive relationships between values plus a trend toward a positive effect of the intervention on utility value. We conclude that content-related interventions in modelling can improve motivational outcomes.

INTRODUCTION

Motivation comprises reasons for human actions (Middleton et al., 2016) and can be seen as a vehicle for human behavior. Prior research has analyzed how different aspects of motivation (e.g., values) are related to students' mathematical well-being (Hill & Seah, 2023) or the overall value of mathematics in the context of learning (Eccles & Wigfield, 2020). According to Eccles' expectancy-value theory of motivation (Eccles & Wigfield, 2020), students' mathematics values are very important for their achievement and educational choices. Values can be ascribed to different objects, such as the domain (e.g., mathematics), a topic (e.g., geometry), or a competency (e.g., modelling) (Schukajlow, Rakoczy, et al., 2023). As students who value mathematics take comprehensive mathematics courses in high school and choose mathematics as a study domain at universities, we addressed values in mathematics in this study.

One way to improve mathematical thinking is to teach students how to solve modelling problems. Solving modelling problems requires demanding processes through which information is transferred between the real world and mathematics (Niss & Blum, 2020). Openness is one important characteristic of modelling problems. However, the openness of modelling problems was found to be a source of various barriers that tend to arise in the solution process (Schukajlow, Krawitz, et al., 2023).

The aims of this study were (1) to analyze the relationships between three different types of values (i.e., attainment, intrinsic, and utility values; see below) and (2) to investigate how a teaching intervention aimed at improving students' ability to solve open modelling problems affects values in 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. 4, pp. 81–88). PME.

THEORETICAL BACKGROUND

Task values as motivational outcomes

Task values are an important component of affective traits (Hannula, 2012). Values are stable motivational dispositions that can be related to learning and achievement. They can be distinguished from more temporary variable states, such as experiences of competence or autonomy in specific learning situations (Schukajlow, Rakoczy, et al., 2023). Task values indicate the personal importance of the tasks, such as the value of one's ability to solve a problem, the value of performing a calculation, or the value of making a drawing to solve a modelling problem. Expectancy-value theories assume that different learners ascribe different values to different tasks and thus, task values are subjective (Eccles & Wigfield, 2020). Furthermore, the extents to which one person values tasks vary across different tasks and different learning situations, indicating the situated nature of task values. In Eccles' expectancy-value theory, Eccles and colleagues proposed three key components of task values: attainment value, intrinsic/interest value, and utility/extrinsic value (Eccles & Wigfield, 2020). If a student sees mathematics as a part of their personality, ascribes mathematical achievement high personal relevance, and strongly identifies with mathematics, the student ascribes high attainment value to mathematics. The intrinsic value of mathematics concurs with enjoyment in solving mathematical problems and enjoyment in engaging in mathematical activities. The utility value of mathematics is reflected in the importance of mathematics for present or future plans, such as the importance of mathematics for learning in school, school grades, career, future work opportunities, or everyday life. Empirical studies have indicated that task values are positively related but distinct factors (Eccles & Wigfield, 2020). However, many studies have addressed the overall task value level by aggregating attainment, intrinsic, and utility values into one score or by using items that referred to the overall value of the presented mathematical problems (Böswald & Schukajlow, 2023; Rach, 2023). Very few studies have analyzed attainment, intrinsic, and utility values as distinct factors in mathematics and offered a more differentiated picture of the development of these motivational outcomes. One exception is a study by Gaspard et al. (2015). In this study, the authors found positive relationships (Pearson's correlations ranged from .50 to .71) between the attainment, intrinsic, and utility values of lower secondary school students. Furthermore, not many studies have analyzed how to improve values in mathematics. One potential way to improve students' values is to teach students to solve open modelling problems. But how can this type of problem be described?

Open modelling problems

In the real world, many problems are open, and their solutions require assumptions to be made. Models of problem solving for open problems distinguish between the openness of the initial state, the openness of transformation, and the openness of the goal state. In modelling problems, the transformation is open because of the need to construct a mathematical model and select appropriate mathematical procedures to solve a problem. Depending on the type of open modelling problem, either the initial state or a goal state can be open (Schukajlow, Krawitz, et al., 2023). Problems with an open initial state do not include all the information needed for their solution. In problems with open goal states, the question is ambiguous, requiring interpretations about the quantity to be calculated to solve the problem. We know from prior research that dealing with openness is demanding for students and pre-service teachers (Galbraith & Stillman, 2001; Stylianides & Stylianides, 2023). In the current study, we focused on problems with an open initial state and a closed goal state. Analyses of cognitive demands that students face while solving modelling problems with an open initial state have indicated that noticing openness, identifying missing quantities, and making realistic assumptions about the missing quantities are essential prerequisites for processing the problems (Schukajlow, Krawitz, et al., 2023). For example, in the "Poster" problem, among other aspects, students should notice that the information about the diameter of the poster roll is missing, and they must make an assumption about its length. The goal state is closed in this task because the goal of the problem is to find out whether the poster will fit in the suitcase. To achieve this goal and solve the problem, students should compare the measurements of the poster roll with those of the suitcase. To do so, they need to calculate the diagonal of the interior of the suitcase.

Poster

Sandy is on vacation in Japan and would like to buy a movie poster there and roll it up to take home in her suitcase. However, she is unsure about whether it is possible to fit the poster in her suitcase. In a store, she finds a poster for 1075 yen. The poster is 105 cm long and 75 cm wide, and Sandy's suitcase is 40 cm long, 25 cm wide, and 60 cm high. When she rolls the longer side of the poster up, she gets a roll that is 75 cm long.

Can she transport the rolled-up poster in her suitcase?

Figure 1: The open modelling problem "Poster"

Interventions to improve task values

Findings from prior research on the effects of interventions that were aimed at increasing students' task values in different domains (e.g., mathematics) and specifically on the task values of solving modelling problems are mixed. Several studies have revealed that emphasizing the relevance of mathematics by prompting students to reflect on it affected task values and related outcomes (Rosenzweig et al., 2022; Schukajlow, Rakoczy, et al., 2023). Furthermore, the types of problems offered to students were shown to affect students' values. In a study of university students, future teachers reported higher intrinsic and utility values regarding profession-related tasks than for tasks that were not related to the teaching of mathematics in schools (Rach & Schukajlow, 2023). Because of the strong relationship between open modelling problems and the real world, students might value this type of problem more than problems that are not that strongly related to reality. However, contrary to this theoretical consideration, students and pre-service teachers did not assign higher overall value to solving open modelling problems compared with solving word or intra-mathematical problems (Böswald & Schukajlow, 2023). Two explanations for these findings are the higher perceived difficulty of open modelling problems and students' lack of familiarity with this type of problem. Teaching students how to deal with openness and how to solve open problems can increase the utility value of open problems and more generally the utility value of mathematics. As engagement in solving open modelling problems will promote the utility value of mathematics, and because of the positive relationships between utility value and the other values, we also expected an increase in attainment and intrinsic values.

RESEARCH QUESTIONS AND EXPECTATIONS

This study was carried out in the framework of the Open Modelling Problems in Self-Regulated Teaching (OModA) project. In this project, we have been investigating how students solve open modelling problems and how teaching can support students' learning of open modelling problems and improve affective outcomes, including the extent to which students value mathematics (Schukajlow, Krawitz, et al., 2023).

Building on expectancy-value theory and the theory on modelling, the research questions in this study were:

RQ1: Are attainment, intrinsic, and utility values in mathematics positively related before the intervention?

On the basis of expectancy-value theory (Eccles & Wigfield, 2020) and prior empirical findings (Gaspard et al., 2015), we expected positive relationships between the three types of values.

RQ2: How does teaching students to solve open problems affect their attainment, intrinsic, and utility values?

We expected that teaching students how to solve open problems (compared with closed real-world problems) would increase students' feelings that mathematics is a part of their personality (attainment value), their enjoyment of mathematics (intrinsic value), and their perceptions that mathematics is useful (utility value).

METHODOLOGY

One hundred eighty-five ninth graders from German middle and high track schools participated in this study (103 female; 14.5 years of age). Within each class, the students were randomly assigned to one of two groups. Students in the experimental group (EG) were taught how to solve open modelling problems, and students in the control group (CG) were taught how to solve closed real-world problems. Before and after the teaching intervention, students filled out questionnaires on values.

The EG learned how to deal with openness, missing quantities, and assumptions for solving open modelling problems. In the CG, students focused on how to solve problems with superfluous information. The problems in the CG did not require students to make assumptions about missing information in order to solve the problems. In both conditions, the teaching unit took 4 x 45 minutes. The study was conducted during regular classes in schools. Within the EG and the CG, students worked in smaller groups to solve the problems and then reflected on their solutions with the whole group (EG or CG) at the end of the class.

Pre-service teachers with bachelor's degrees in mathematics education served as teachers in this study. Before the study, they were given standardized training. To balance the effects of the instructor's personality on students, each teacher taught students in both EGs and CGs.

To assess values, we used well-evaluated Likert scales (ranging from 1 = not at all true to 5 = completely true) from prior studies. The scales on attainment, intrinsic, and utility values included 3 items each, and the scales' internal consistencies (Cronbach's Alpha) were at least acceptable (i.e., higher than .78). Sample items from the attainment, intrinsic, and utility value scales are: "It is important for me to be a person who is good at mathematics" (attainment value), "I like mathematics" (intrinsic value), "Mathematics is useful for my future life" (utility value).

To check the fidelity of the treatment, student assistants observed how the EGs and CGs were taught. Student assistants used a standardized observation questionnaire in which they were asked to note any deviations from the instructional manuals. In addition, we collected all student materials so that we could analyze the treatment fidelity. The analysis of the observational questionnaires and teaching materials indicated high treatment fidelity. In all classes, the teachers gave the tasks to the students in the same order in the EG and CG, and the teachers followed the teaching manual closely.

To address the research questions, we calculated Pearson correlations and conducted repeated-measures ANOVAs. In our statistical analysis, we included students who participated in the intervention (in the EG or CG) and filled out questionnaires at pretest and posttest ($N = 185$). Students who missed the intervention or one of the tests were excluded from the analysis. The percentage of missing values ranged from 16.6% for utility value on the pretest to 21% for utility value on the posttest.

RESULTS

The analysis of the correlations between the three values at pretest was in line with our expectations. Attainment value was positively related to intrinsic value ($r = .60$, $p < .001$) and utility value ($r = .54$, $p < .001$), and intrinsic value was positively related to utility value ($r = .51$, $p < .001$).

We conducted three repeated-measures ANOVAs with time as a within-subject factor (pretest, posttest) and treatment as a between-subject factor (EG, CG) and attainment value, intrinsic value, or utility value as dependent variables. The statistical analyses revealed mixed results. Contrary to our expectations, we did not find an effect of the interaction between the time and treatment factors on attainment value ($M_{EG,\ pre}\ (SD) =$

2.89 (0.93), $M_{CG,\ pre}\,(SD)$ = 2.82 (0.95); $M_{EG,\ post}\,(SD)$ = 2.77 (1.03), $M_{CG,\ post}\,(SD)$ = 2.59 (1.09); time*treatment: $F(1, 183) = 1125$, $p = .14$, $\eta^2 = 0.006$) or on intrinsic value ($M_{EG,\ pre}\,(SD)$ = 2.67 (1.03), $M_{CG,\ pre}\,(SD)$ = 2.63 (1.08); $M_{EG,\ post}\,(SD)$ = 2.43 (1.01), $M_{CG,\ post}\,(SD)$ = 2.50 (1.11); time*treatment: $F(1, 183) = 1093$, $p = .15$, $\eta^2 = 0.006$). These results indicate that teaching students how to solve open problems did not positively affect their attainment value or intrinsic value. The analysis of the effects on utility value indicated different results. In line with our expectations, the effect of the intervention on utility value was significant at the 10% level ($M_{EG,\ pre}\,(SD)$ = 3.23 (0.93), $M_{CG,\ pre}\,(SD)$ = 2.28 (0.94); $M_{EG,\ post}\,(SD)$ = 3.27 (0.90), $M_{CG,\ post}\,(SD)$ = 3.17 (1.01); time*treatment: $F(1, 183) = 1093$, $p = .09$, $\eta^2 = 0.009$). Thus, students who were taught how to solve open modelling problems tended to benefit more than students who were taught to solve closed real-world problems with respect to utility value.

DISCUSSION

The analysis of the relationships between the three values in mathematics indicated positive relationships, supporting theoretical considerations from expectancy-value theory and results from prior studies (Eccles & Wigfield, 2020; Gaspard et al., 2015). One practical implication from this study is that it might be possible to primarily address one of the values in a teaching intervention, such as utility value, because the other values may then be affected through utility value.

To analyze how teaching students to solve open modelling problems affects students' motivation, we set up a randomized quasi-experimental study. Because of the significance of task values for achievement and educational choices, we selected students' values as motivational outcomes that can be affected by teaching. On the basis of expectancy-value theory (Eccles & Wigfield, 2020) and considerations from research on modelling (Niss & Blum, 2020; Schukajlow, Krawitz, et al., 2023), we hypothesized that if students learned how to solve open modelling problems (i.e., problems that are closely related to their real lives), they might subsequently value mathematics to a greater extent. After the teaching intervention, changes in students' attainment, intrinsic, and utility values were expected to be more beneficial for students in the EG that for students in the CG, who solved closed real-world problems. In line with our expectations, we found a positive trend toward an effect of teaching students to solve open modelling problems on utility value. This finding means that students who learned how to solve open problems tended to report higher utility value compared with students who learned how to solve closed real-world problems. This result supports theoretical considerations that, while learning how to solve open modelling problems, students perceive the relevance of mathematics for real life, and thus, they might understand how useful mathematics can be for their future lives or careers. This important result is in line with studies that have demonstrated the positive effects of interventions on the relevance of mathematics on values (Rosenzweig et al., 2022). Furthermore, this result supports the importance of the types of problems that students deal with in mathematics classes for the development of utility value. This consideration is hypothesized in expectancy-value theory and received empirical support in a prior study (Rach & Schukajlow, 2023). The theoretical implication of this study within the framework of expectancy-value theory is the indication that it may be possible to increase the fit between students' personal plans and doing mathematics by teaching open modelling problems. Future studies should clarify whether teaching open modelling problems fit into students' already existing plans (i.e., to use mathematics in the real world) or whether this intervention helped students develop such plans (i.e., helped them see the usefulness of mathematics in the real world).

Another important result of our study is the lack of effects of teaching students to solve open problems on attainment and intrinsic values in mathematics. One explanation might lie in how values in mathematics were assessed. Using modelling competence as an object of the values might reveal other results. The alignment between the object of the intervention and the measures may be a significant factor that affected the results of the study. In future studies, researchers should assess values with respect to open modelling problems and closed word problems (see an example of assessment in Böswald & Schukajlow, 2023). Another explanation might be that making mathematics personally important and improving enjoyment in mathematics might require more comprehensive instruction. Indeed, prior interventions have rarely targeted these values in the past (Rosenzweig et al., 2022), and it is very important to develop ideas about how teaching modelling problems can specifically improve attainment and intrinsic values. One way to increase intrinsic value might be to use contexts that refer to students' everyday lives (e.g., cities where students live), to adjust the context to students individual values (Bernacki & Walkington, 2018) or to ask them to pose their own problems (Voica et al., 2020). Furthermore, it is important to clarify in future studies what cognitive and motivational states during instruction mediate the effects of teaching students how to solve modelling problems on their task values.

Acknowledgments: This study was financially supported by the German Research Foundation (GZs: RA 1940/2-1 and SCHU 2629/5-1).

References

Bernacki, M. L., & Walkington, C. (2018). The role of situational interest in personalized learning. *Journal of Educational Psychology*, *110*(6), 864–881. https://doi.org/10.1037/edu0000250

Böswald, V., & Schukajlow, S. (2023). I value the problem, but I don't think my students will: Preservice teachers' judgments of value and self-efficacy for modelling, word, and intramathematical problems. *ZDM - Mathematics Education*, *55*, 331–344. https://doi.org/10.1007/s11858-022-01412-z

Eccles, J. S., & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology*, *61*, Article 101859. https://doi.org/10.1016/j.cedpsych.2020.101859

Galbraith, P. L., & Stillman, G. (2001). Assumptions and context: Pursuing their role in modelling activity. In J. Matos, W. Blum, K. Houston, & S. Carreira (Eds.), *Modelling and Mathematics Education, Ictma 9: Applications in Science and Technology* (pp. 300–310). Horwood Publishing. https://doi.org/10.1533/9780857099655.5.300

Gaspard, H., Dicke, A.-L., Flunger, B., Brisson, B. M., Häfner, I., Nagengast, B., & Trautwein, U. (2015). Fostering adolescents' value beliefs for mathematics with a relevance intervention in the classroom. *Developmental psychology*, *51*(9), 1226–1240. https://doi.org/10.1037/dev0000028

Hannula, M. S. (2012). Looking at the third wave from the West: framing values within a broader scope of affective traits. *ZDM Mathematics Education*, *44*(1), 83–90. https://doi.org/10.1007/s11858-012-0410-5

Hill, J. L., & Seah, W. T. (2023). Student values and wellbeing in mathematics education: perspectives of Chinese primary students. *ZDM - Mathematics Education*, *55*, 385–398. https://doi.org/10.1007/s11858-022-01418-7

Middleton, J. A., Jansen, A., & Goldin, G. A. (2016). Motivation. In G. A. Goldin, M. S. Hannula, E. Heyd-Metzuyanim, A. Jansen, R. Kaasila, S. Lutovac, P. Di Martino, F. Morselli, J. A. Middleton, M. Pantziara, & Q. Zhang (Eds.), *Attitudes, Beliefs, Motivation, and Identity in Mathematics Education* (pp. 17–22). Springer.

Niss, M., & Blum, W. (2020). *The learning and teaching of mathematical modelling*. Routledge.

Rach, S. (2023). Motivational states in an undergraduate mathematics course: relations between facets of individual interest, task values, basic needs, and effort. *ZDM – Mathematics Education*, *55*, 461–476. https://doi.org/10.1007/s11858-022-01406-x

Rach, S., & Schukajlow, S. (2023). Affecting task values, costs, and effort in university mathematics courses: the role of profession-related tasks on motivational and behavioral states. *International Journal of Science and Mathematics Education*. https://doi.org/10.1007/s10763-023-10413-7

Rosenzweig, E. Q., Wigfield, A., & Eccles, J. S. (2022). Beyond utility value interventions: The why, when, and how for next steps in expectancy-value intervention research. *Educational Psychologist*, *57*(1), 11–30. https://doi.org/10.1080/00461520.2021.1984242

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. https://doi.org/10.1007/s10649-023-10265-6

Schukajlow, S., Rakoczy, K., & Pekrun, R. (2023). Emotions and motivation in mathematics education: Where we are today and where we need to go. *ZDM - Mathematics Education*, *55*, 249–267. https://doi.org/10.1007/s11858-022-01463-2

Stylianides, G. J., & Stylianides, A. J. (2023). Promoting elements of mathematical knowledge for teaching related to the notion of assumptions. *Mathematical Thinking and Learning*, 1–29. https://doi.org/10.1080/10986065.2023.2172617

Voica, C., Singer, F. M., & Stan, E. (2020). How are motivation and self-efficacy interacting in problem-solving and problem-posing? *Educational Studies in Mathematics*, *105*(3), 487–517. https://doi.org/10.1007/s10649-020-10005-0

SIMULATIONS OF PROBLEM-BASED LESSONS: USING A CONJECTURE MAP TO RELATE DESIGN AND OUTCOMES

Gil Schwarts, Patricio Herbst, and Amanda Brown

University of Michigan, USA

Virtual simulations are a promising tool for mathematics teachers' preparation and professional development. This paper focuses on a set of simulations of a problem-based lesson, illustrating how their design informed prospective teachers' changes in decision-making during simulations. Using a design-based approach and conjecture mapping, we trace the processes that could be attributed to the observed changes in teachers' decisions. The analysis shows that after completing the set of simulations, teachers increased their selection of student work that relates to the lesson goal, regardless of its correctness. The paper contributes to the understanding of virtual simulations as a sustainable tool for teacher preparation and professional learning.

BACKGROUND AND AIMS

Practice-based virtual simulations are a sustainable resource for teacher education, leveraging emerging technologies to provide mathematics teachers with immersive occasions for approximating practice (e.g., Mikeska et al., 2023). This paper explores the design of an online, self-paced model for mathematics teachers' education and professional development (PD) that consists of digital simulations of problem-based lessons where secondary mathematics teachers make decisions for a teacher avatar.
The simulations are aimed at supporting teachers in problem-based instruction, that is, in teaching lessons that revolve around a novel problem on which students work to arrive at a new curricular goal. A main role for the teacher in such lessons is to leverage students' genuine ideas to reach the lesson goal. This includes tasks like selecting and sequencing student work (hereafter SW) and orchestrating whole-class discussions around these pieces of work (see the "five practices" model; Stein et al., 2008).

The interactive nature of problem-based instruction makes it challenging for teachers to prepare for such lessons. Virtual simulations present a valuable means for aiding teachers growing in this regard, providing them with an environment of reduced complexity where they can practice teaching with avatars (e.g., Dieker et al., 2014). However, the processes for achieving such growth are oftentimes implicit. In particular, empirical research that specifies how the design of virtual teacher simulation informs teacher learning is still emerging.

A critical aspect of simulation design revolves around determining the type and timing of feedback. Simulations may offer immediate feedback, allowing teachers to adapt their decision-making in real-time, but this might reduce their presence. Alternatively, feedback can be provided after the fact. The nature of feedback varies significantly, as evidenced by two U.S.-based simulation models, simSchool (Tyler-Wood et al., 2015)

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. 4, pp. 89–96). PME.

and TeachLive/Mursion (Dieker et al., 2014). SimSchool offers *didactic* feedback by assessing the teacher's actions against predetermined criteria for appropriateness. In TeachLive, simulations involve behind-the-scenes inter-actors playing students, providing *adidactic* ("soft") feedback through the reactions of the simulated students to the teacher's decisions. The latter model prompts the question: in the absence of peers, facilitators, and *didactic* feedback, how do teachers learn in virtual simulations?

We address this overarching question by presenting the design of pre-programmed simulations (without actors) utilizing storyboards to depict classroom scenarios. The model relies on adidactical feedback, implemented through visualizations of contingencies tied to participants' choices for a teacher avatar. To discuss this design and share preliminary results from a cycle of enactment with prospective teachers (PTs), we follow a design-based research paradigm. Our focal research question is:

How and what do PTs learn about problem-based instruction from participating in a set of adidactical, online, self-paced simulations of a problem-based lesson?

THEORETICAL FRAMEWORK

The design of the represented problem-based lesson (including dialogues and artifacts) was guided by key themes from *practical rationality* (Herbst & Chazan, 2011). In this framework, *instructional situations* are recurrent types of tasks which can be described by subject-specific norms. In U.S. geometry lessons, examples of different instructional situations are *proof*, *construction*, and *calculation*, and teachers can either *frame* a problem as a case of an instructional situation, or leave it vague (for example, consider the difference between asking students to *construct* a circle, versus *finding* one). Herbst et al. (2023) argue that situational norms and the lesson goal are helpful for describing teachers' categories of perception when engaging with student ideas. The category of *normativity* alludes to a teacher's perception of how closely a student idea aligns with the norms of the instructional situation used to frame the problem. For example, if a problem is framed as one expecting students to *construct*, teachers may consider a sketched diagram as less normative than a diagram for which construction tools were used. Crucially, normativity differs from correctness; it pertains to adherence to the norms of the instructional situation, meaning a normative SW might be either correct or incorrect (see Figure 3b for an incorrect normative SW). A second category of perception is *serviceability*, which attends to the teacher's perception of the usefulness of a student's contribution to reach the lesson content goal. In this regard, a sketch that introduced ideas relevant to the lesson goal might be considered more serviceable, even if it is less normative or incorrect, compared to a construction that cannot be leveraged for arriving at the lesson goal (see Figure 3a). Overall, these two categories of perception allow researchers to describe the work of mathematics teachers handling student work. They were used in our simulation design.

METHODS

Design-based research and conjecture mapping

Design-based research (DBR) aims to pursue both practical improvement and theoretical refinement (Sandoval, 2014, p. 19), through cycles of design, enactment, analysis, and revision. To provide adequate argumentative grammar (Kelly, 2004) for DBR, *conjecture mapping* helps explicate the links between design elements, theoretical conjectures, and observed outcomes (Sandoval, 2014). Conjecture maps consist of the following components: (1) *High-level conjecture*: a general statement about how an intervention intends to support some form of learning; (2) *Embodiment*: the physical artifacts used (e.g., instruments, software, media), the task structure (what participants are asked to do) and the social structure (how participants are expected to relate to each other during the task); (3) *Mediating processes*: the processes in which the design features are supposed to activate learning (assuming that the design itself does not directly produce specific outcomes); and (4) *Desired outcomes*: the expected results of the mediating processes, that should be clearly stated for easy measurement. In this paper we elaborate on one cycle of our simulation project, showing how the enactment phase relates to the conjecture map and informs its revisions.

Participants and data analysis

Our analysis is based on implementing the simulations with 11 PTs from the Northwest U.S. (5 male, 6 female). Over a 5-week period as part of their methods course in mathematics education, they individually engaged with the simulation set. The PTs had opportunities to discuss their decisions in class. The analyzed data comprise the decisions made by PTs in the simulations (including both open-ended and close-ended items), along with our design documents that informed the conjecture map below. We employed mixed methods to discern changes in PTs' decisions: For closed-ended items involving the selection and sequencing of student work, we created scores for correctness, normativity, and serviceability (details in the embodiment section below). Having pre- and post-data on participants' decisions, we employed a two-tailed Wilcoxon test to assess the significance of the observed changes. The open-ended items (e.g., that asked for justifications for decisions) were subject to content analysis guided by the same categories, while also considering emergent themes. By linking scores and codes to the design elements, we could articulate and test the design conjectures.

DESCRIPTION OF INTERVENTION AND PRELIMINARY FINDINGS

The following section elaborates on the conjecture map's components (Figure 1).

High-level conjecture. The design of the simulation was part of a broader project focused on studying how teachers make decisions in problem-based lessons, and enhancing their capacity to teach such lessons in ways that involve student voice. Within this project, the simulations build on a collaborative professional learning model – Story*Circles* (Herbst & Milewski, 2018) – where teachers co-design problem-based lessons. Our aim in adapting Story*Circles* into virtual simulations was to create a scalable, on-demand PD model that does not rely on peers and facilitators. In this context, the high-level conjecture is that simulations that provide teachers with the opportunity to practice teaching a problem-based lesson that builds on student ideas can increase teachers' capacity to anticipate and manage such a lesson.

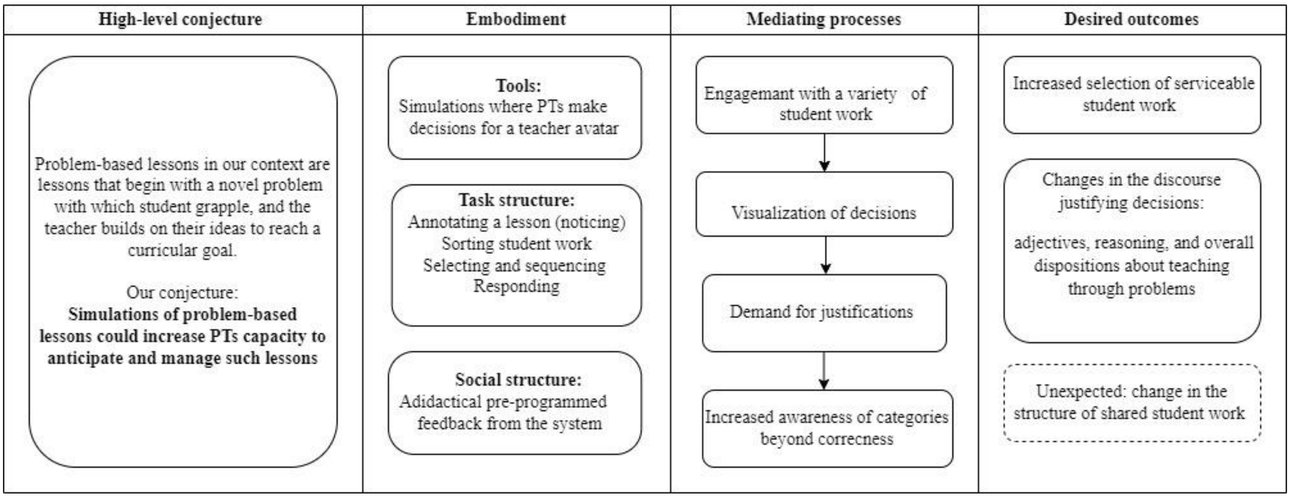


Figure 1: A conjecture map for the set of simulations

Embodiment. We designed four simulations of a lesson centered on a problem that asks students to find a circle tangent to two intersecting lines, ultimately leading to the tangent segments theorem (*Two intersecting lines are tangent to a circle if and only if the points of tangency are equidistant from the point of intersection of the lines*). Two simulations covered the entire lesson: a) "Getting to know the lesson," where participants were introduced to the problem, through a teacher-centered lesson that discussed possible solutions to the problem and the lesson goal (the aforementioned theorem) and were tasked with annotating the lesson's representation; and b) "Teaching the lesson with student participation," where participants made decisions for a teacher avatar in key moments of the lesson (e.g., framing the problem, selecting and sequencing SW, responding to students at the board). The other two simulations targeted specific phases in the lesson: c) "Anticipating student work" focuses on the phase when students grapple with the problem, and the teacher monitors their work, and in d) "Responding to student work," participants saw storyboarded scenes where students shared their work publicly and were asked about their goals for handling the student contribution. In more detail, in simulation (c), participants viewed a variety of SW, sorted the samples into bins (see Figure 2), named the bins, and explained their sorting to hypothetical colleagues. Participants then selected and sequenced SW. The SW samples were carefully designed to feature variations in correctness, normativity (alignment with the instructional situation of construction – e.g., using tools), and serviceability (usefulness for arriving at the theorem, e.g., drawing perpendiculars and locating the circle's center in their intersection). Respectively, the SW samples were apriori coded with "1" and "0" (e.g., Sigma's work in Figure 3a was coded [1,0,1] because it is correct, non-normative, and serviceable, while Phi's work in Figure 3b was coded [0,1,0] because it is incorrect, normative, and non-serviceable). As mentioned, the SW samples exhibited a variety of combinations of these codes. In simulation (d), participants viewed a collection of responding moves that attended to serviceable, normative, or generic aspects of student ideas. Participants were offered a subset of recommended moves that relate to speech functions they indicated (e.g., open/close; support/confront; Milewski & Strickland, 2020) and were asked to choose from the proposed moves.

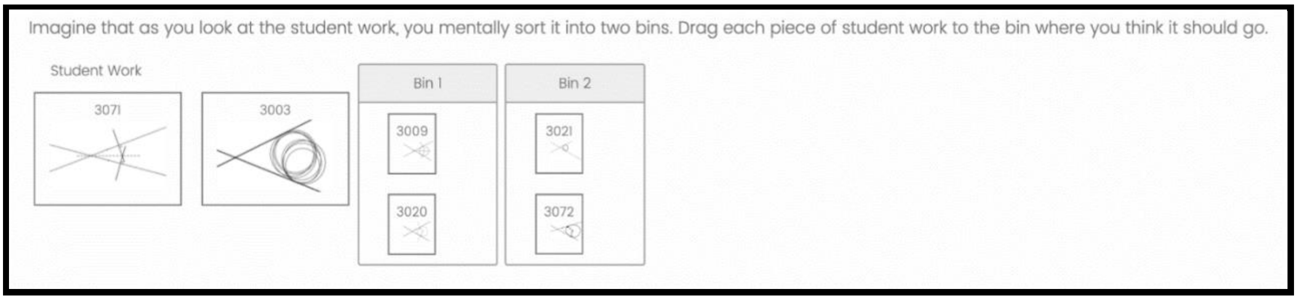


Figure 2: A sorting activity in the Anticipating Student Work simulation

While in simulation (a) participants could only annotate the lesson, simulations (b), (c), and (d) incorporate pre-programmed branches that respond to participants' choices (in a "choose your own adventure" style). To achieve this, many decisions are closed-ended items. However, the simulations also include numerous open-ended items, prompting participants to write notes about student work, elucidate their rationale for choices, suggest teaching moves, anticipate students' responses, and more. In terms of participant structure, in contrast with other models of PD, the simulations engage teachers individually. Nevertheless, as outlined above, the system offers various forms of feedback, including showcasing how other teachers addressed the same decisions and providing opportunities to revise choices.

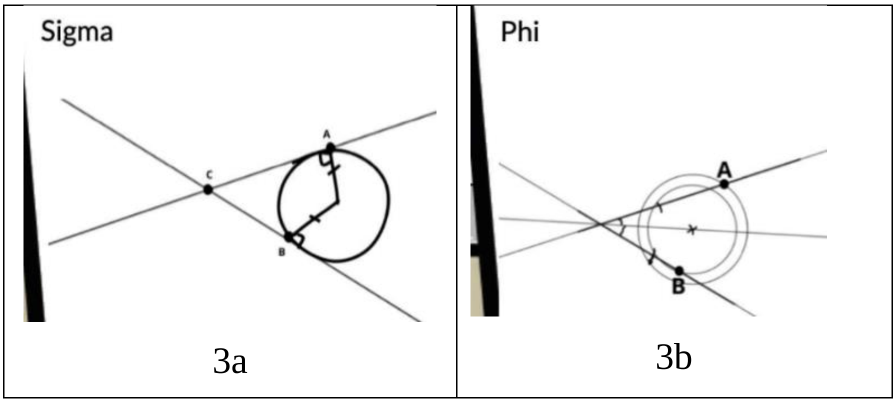


Figure 3: Samples of student work that were used in the simulations

To understand how simulations (c) and (d) could influence the decisions teachers make during a complete lesson, participants underwent simulation (b) twice: first in the second week (following simulation (a)) and then in the final week (after simulation (d)). In summary, the sequence of simulation engagement was: a, b, c, d, b.

Mediating processes. The meticulous design of each simulation intends to activate mediating processes in varying grain-sizes, some are immediate (e.g., sorting student work into 3-5 bins instead of just 2—intended to foster awareness of categories beyond correct/incorrect), while others are more complex. Due to space constraints, we elaborate on one process: participants' engagement with samples of student work. Our mid-level conjecture was that initially, *participants would prefer correct and normative student work (even when not serviceable) over samples that are serviceable but lack correctness or normativity*. The associated tasks include sorting various student work into bins (Figure 2), naming bins, observing how other teachers have sorted it, selecting and sequencing SW, and justifying how these decisions align with the lesson content goal. The mediating processes, thus, involve being exposed to a variety of student work, the visualization of the consequences of decisions about SW, and the demand for justifications of decisions, which together are supposed to lead to increased awareness of: 1) categories of perception beyond correctness – specifically, links between selected SW and the lesson goal; and 2) the idea that an informed selection of SW could leverage the ensuing discussion.

Desired and observed outcomes. The desired outcome from these processes is an enhanced focus on serviceability among participants, manifested in an increased selection of serviceable student work and changes in the discourse justifying decisions (e.g., changes in adjectives describing student work, reasoning for selection, and overall dispositions about teaching through problem-solving). Our preliminary analysis indicates the achievement of some of these outcomes. Using the scores outlined in the Methods section, we conducted a two-tailed Wilcoxon test to compare participants' decisions on selecting student work in the second and fifth simulations (which were identical). The results indicate that, while normativity and correctness scores did not significantly change, the serviceability score increased significantly (p-value = 0.026, Z = -2.232). These findings are supported by the content analysis of the justifications provided. In the initial round, participants cited reasons for selecting and sequencing such as:

DK: Start with one that has common errors then roll into one that adjusted their approach to talk about the relations it made and then end with one that's correct.

GM: I want to start with clearing any misconceptions.

In the second round, the emphasis changed:

MZ: None of them are correct but I am hoping they notice something about the points that they are choosing for B.

ZI: […] I then want to introduce Sigma's work because although it is not precise, I'm hoping the class can still learn from their idea that the tangent line is perpendicular to the radius.

A surprising finding was that some participants shifted from a conception of sequencing that gradually approximates the correct solution (ZS: "it is important to see the progression of detail") to selecting student work that features complementary aspects of the solution (ZS: "They are each bringing something different to the table") or, alternatively, showcasing different ways to solve. We interpret this change as additional evidence of participants' heightened awareness of student epistemic agency: the latter approach offers more opportunities for generative connections between samples of SW. In addition to these findings, the analysis revealed that while in the first round the most frequently chosen summary statement was procedural and detailed, in the second round, the prevalent statement discussed the solvability of the problem but did not provide students with instructions on how to continue.

CONCLUDING REMARKS

This paper describes the processes by which a set of virtual simulations support PTs' learning about problem-based instruction. This is crucial as the mechanisms of teacher learning in virtual simulations are not yet well specified. In addition, we shed light on a specific kind of teacher learning that promotes teacher agency. This approach is manifested in a key aspect of our design – providing mostly adidactical feedback. In choosing such a design, our approach for teacher education coheres with the disposition that students (here, PTs) should have chances to develop epistemic agency, rather than depend on an outside source of knowledge.

The conjecture map assisted us in clarifying hypotheses, identifying relationships, and distinguishing between expected and unexpected outcomes (Sandoval, 2014). The unexpected outcomes also give rise to the identification of emerging mediating processes – for example, the idea that sorting SW into bins should support the selection of SW – thus informing the revisions of the current map and of the simulation design. While the conjecture map helped explicate delicate processes, it has limitations, such as our lack of hard evidence that the mediating processes are indeed the reason for the observed changes. Moreover, we lack details on the discussions about features of student work in the PTs' methods course, and thus the course is not part of the map. The small sample size is another limitation. Nevertheless, this limitation allowed us to generate a blueprint for the desired outcomes with a larger sample, data that we are currently collecting.

In summary, this work contributes to understanding technology-mediated, self-paced teacher learning, offering a novel approach to enhance teachers' readiness for attending to and building on diverse sets of student ideas.

Additional information

Research supported by James S. McDonnell Foundation grant 220020524.

References

Dieker, L. A., Rodriguez, J. A., Lignugaris/Kraft, B., Hynes, M. C., & Hughes, C. E. (2014). The potential of simulated environments in teacher education: Current and future possibilities. *Teacher Education and Special Education*, *37*(1), 21–33. https://doi.org/10.1177/0888406413512683

Herbst, P., Brown, A. M., Chazan, D., Boileau, N., & Stevens, I. (2023). Framing, responsiveness, serviceability, and normativity: Categories of perception teachers use to relate to students' mathematical work in problem-based lessons. *School Science and Mathematics*, *123*(7), 398–413. https://doi.org/10.1111/ssm.12600

Herbst, P., & Chazan, D. (2011). Research on practical rationality: Studying the justification of actions in mathematics teaching. *The Mathematics Enthusiast*, *8*(3), 405–462. https://doi.org/10.54870/1551-3440.1225

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

Kelly, A. E. (2004). Design research in education: Yes, but is it methodological? *Journal of the Learning Sciences*, *13*(1), 115–128. https://doi.org/10.1207/s15327809jls1301_6

Mikeska, J. N., Howell, H., & Kinsey, D. (2023). Do Simulated Teaching Experiences Impact Elementary Preservice Teachers' Ability to Facilitate Argumentation-Focused Discussions in Mathematics and Science? *Journal of Teacher Education*, *74*(5), 422–436. https://doi.org/10.1177/00224871221142842

Milewski, A. M., & Strickland, S. K. (2020). Building on the work of teachers: Augmenting a functional lens to a teacher-generated framework for describing the instructional practices of responding. *Linguistics and Education*, *57*, 100816. https://doi.org/10.1016/j.linged.2020.100816

Sandoval, W. (2014). Conjecture mapping: An approach to systematic educational design research. *Journal of the Learning Sciences*, *23*(1), 18–36. https://doi.org/10.1080/10508406.2013.778204

Stein, M. K., Engle, R. A., Smith, M. S., & Hughes, E. K. (2008). Orchestrating productive mathematical discussions: Five practices for helping teachers move beyond show and tell. *Mathematical thinking and learning*, *10*(4), 313–340. https://doi.org/10.1080/10986060802229675

Tyler-Wood, T., Estes, M., Christensen, R., Knezek, G., & Gibson, D. (2015). SimSchool: An opportunity for using serious gaming for training teachers in rural areas. *Rural Special Education Quarterly*, *34*(3), 17–20. https://doi.org/10.1177/875687051503400304

A PRELIMINARY ANALYSIS OF TWO PROOF LESSONS FROM AN INTERNATIONAL COMPARATIVE PERSPECTIVE: A CASE STUDY ON GERMAN AND JAPANESE GRADE 8 CLASSROOMS

Yusuke Shinno[1], Fiene Bredow[2], Christine Knipping[2], Ryoto Hakamata[3], Takeshi Miyakawa[4], Hiroki Otani[5], and David Reid[6]

Hiroshima University[1], University of Bremen[2], Kochi University[3], Waseda University[4], Otsuma Women's University[5], and University of Agder[6]

In this paper, we aim to offer a preliminary analysis of two proof lessons from a comparative perspective. A case study focuses on German and Japanese grade 8 classrooms, where a common topic for a proof in algebra is introduced. Adapting Boero's (1999) six phases, we discussed how the two lessons are organized differently in 'pre-activities', 'exploration of proof ideas', and 'formulation of a proof'.

INTRODUCTION

Proof and proving is seen as a well-developed but still developing fields of research in mathematics education. Recently, researchers in this field have paid more attention to studying the reality of the classroom (e.g., Mariotti et al., 2018; Stylianides et al., 2023). Another important issue in this field is that more research from an international perspective is needed to prevent the scientific studies from only being useful in a single national or cultural context (Reid et al., 2022).

Our primary concern in this research report is to understand how proof and proving are organized in an ordinary mathematics classroom. From an international perspective, we also investigate and compare the organizations of the teaching of proof through a case study in German and Japanese lessons. Some useful knowledge about teaching mathematics in both countries is available from the TIMSS video study (Stigler & Hiebert, 1999). According to their findings, the teaching patterns of the two countries exhibit contrasting approaches: the German pattern is called *developing advanced procedures* and the Japanese pattern is called *structured problem solving*. Therefore, we expected the organization of teaching proof to be different. In this paper, we aim to offer a preliminary analysis of two proof lessons from a comparative perspective. For this analysis, we use a framework consisting of a set of categories focusing on proof-related phases in a lesson on which a comparison in both countries can be based on.

THEORETICAL PERSPECTIVES

Regarding the theoretical framework of the phases in classrooms, the Theory of Didactical Situations (TDS) (Brousseau, 1997) offers three basic types of situations: *situations of action, situations of formulation*, and *situations of validation*. Although these types of situations can be applied to mathematics lessons in general, we need more focused categories to better understand the context in which proof occurs.

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. 4, pp. 97–104). PME.

Researchers have proposed different theoretical frameworks to characterize the nature of proof. The existing literature suggests that 'proof' is not a stand-alone entity, but it is associated with other aspects. For example, Balacheff (1987) proposed a framework comprising *knowledge*, *formulation*, and *validation*. This framework allows us to understand the complexity of mathematical proof at different levels. Mariotti et al. (1997) proposed a model of the *Mathematical Theorem* consisting of a system of relations between *statement*, its *proof*, and the *theory* within which the proof makes sense. Related to this model, the Italian colleagues also offered a concept of *cognitive unity*, meaning the potential continuity between generation of a conjecture through argumentation and construction of its proof. Furthermore, in relation to the notion of cognitive unity, Boero (1999) conceptualized six phases of mathematical activities concerning theorems: 1) *production of a conjecture*, 2) *formulation of a statement*, 3) *exploration of the content of the conjecture*, 4) *selection and enchaining of arguments into a deductive chain*, 5) *organization of the enchained arguments into a proof*, and 6) *approaching a formal proof*. Adapting Boero's (1999) distinction, we developed a framework consisting of the following seven categories, as listed in Table 1 with their corresponding characterizations and possible relationships with Boero's six phases.

Categories	Descriptions of activities (examples)	Boero (1999)
Pre-activities	Understanding the tasks, working on the examples	-
Exploration of a statement	Discovering a statement (conjecture), understanding the statement	Phase 1
Formulation of a statement	Representing a statement according to shared textual conventions	Phase 2
Exploration of proof ideas	Discovering proof ideas, understanding the proof ideas	Phase 3
Formulation of a proof	Constructing a proof, presenting the proof	Phases 4–6
Reflection	Evaluating a proof, explaining a proof to others, comparing proofs	-
Application activities	Working on similar proving tasks	-

Table 1: Categories for describing proof lesson structure

These categories can be used to describe the structure of a proof lesson and explain classroom phenomena, focusing on different activities regarding a statement and its proof. Our categories allow us to describe a relatively wider moment of activities from educational points of view, while Boero's (1999) distinction allows for a more detailed analysis of the activities in 'formulation of a proof'. As the categories play a descriptive role, rather than a prescriptive one, some categories may not appear, or two (or more) categories may be combined into one category. Additionally, the order of the categories is often organized chronologically, but it depends on the didactical and epistemological necessity (this is similar to the assumptions of didactical situations in the TDS).

The research questions herein are as follows: RQ1) *How are mathematics lessons of proof and proving structured in the cases of Grade 8 classrooms in Germany and Japan?* RQ2) *What are the commonalities and specificities of proof and proving in the two lessons in terms of the categories proposed in this paper?* To answer these questions, we offer a comparative case study but do not claim that the case study is a representative case for all classrooms in each country.

METHODOLOGY

Context and data collection

This study is a part of an international research project on argumentation and proof from a comparative perspective. The case study is conducted by focusing on the lessons that introduce algebraic proofs in German and Japanese Grade 8 classrooms. In the German national standards, proofs are included as a competency (related to 'argumentation'). However, this does not mean that proofs are taught systematically in classrooms, as this depends on different institutional factors (such as school types and provincial curricula). Concepts or theorems related to proofs are usually taught in geometry and little in algebra. In the Japanese national curriculum, proof is a content to be taught in middle schools (Grades 8–9) in both algebra and geometry domains. Proofs and related theorems are taught more systematically in geometry than in algebra domain. There is a curricular constraint in that the term 'proof' should be introduced in Grade 8 geometry. When algebraic proofs (e.g., a proof about the sum of two even numbers) are taught in Grade 8, the term 'explanation' is used instead of 'proof' until the time of teaching geometry proofs.

Data were collected from Grade 8 classrooms with ethical commitments in both countries in 2023: a 'common' school (called *Oberschule*) in Germany and a public ordinary school in Japan. To make the data analysis as comparable as possible, we selected a common topic for both sides: 'the sum of three consecutive numbers (is divisible by three)'. This topic is used as the first introduction to a proof in algebra. The lessons were planned and implemented by the teachers. Consequently, the lessons were designed as 2 consecutive lessons (80 mins in total) for the German case and 1 lesson (50 mins) for the Japanese case. Despite the difference in the durations, we could compare the activities of proof and proving in terms of the lesson structure. The lessons were videotaped and transcribed. For the German case students' worksheets, designed by the teacher, were also collected.

The method of preliminary analysis

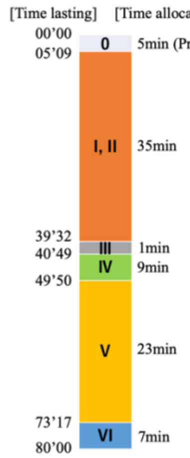
The analytical method of our case study entails a qualitative approach with two steps, which applied to both cases: 1) identifying episodes from the classroom video and the transcripts, and 2) characterizing a lesson structure in terms of the categories. In these steps, we focused on the activities, main tasks, and questions used in the lessons.

Focusing on the analysis in the second step, we first describe the characterization of an overview of the lesson structure for each case, and then explore some aspects of proof and proving according to the categories. These illustrative analyses then allow us to discuss the commonalities and/or specificities in terms of proof lesson structures.

RESULTS

German lesson structure

Figure 1 shows an overview of a German lesson structure consisting of six main phases, which are nearly corresponding to the tasks presented in the worksheet. Although the lesson was basically developed as a form of whole-class instruction, in Phase V, there was a 'differentiation' in this classroom: a group of students (G level/standard) was asked to work on verbal and algebraic proofs by rearranging partial arguments as a whole proof, and another group (E level/higher) was asked to construct an algebraic proof by themselves. We will illustrate some of the activities from Phases I to V.



Phases	Task#	Task descriptions	Categories
I	1	Calculating the following sums. a. 7+8+9, b. 51+52+53, c. 129+130+131, d. 5318+5319+5320	Pre-activities
II	2	Look at the answers and their justifications from the grade five class. a. Which ideas do the kids have? b. Have a look at the justifications and try them out. c. Are all justifications correct? Justify!	
III	-	(No task)	Formulation of a statement
IV	3	Look at the justification of Ayse's statement*. a. Is Ayse's statement correct? (*attributed to a fictional student Ayse) b. Are the kids' justifications correct? Explain, why (not).	Exploration of proof ideas
V	4	(E-level) Prove Ayse's statement using terms and equations. (G-level) a. Bring the proofs into the correct order and fill in the blanks. (G-level) b. Which proof do you prefer? Justify!	Formulation of a proof
VI	-	(Working on a new statement: the sum of five consecutive numbers is divisible by five)	Application activities

Figure 1: An overview of a German lesson structure

Phases I and II (Pre-activities). The lesson began with calculating tasks (such as '7+8+9', '129+130+131') and asking students to find 'smart' strategies for calculations. In Phase II, three strategies were given in the worksheets and students discussed if the strategies worked correctly. For example, one of the strategies, 'this is 3 times the number in the middle', is explained by the operations such as '(129+1)+130+(131-1)=130+130+130=3*130=390'. The teacher wrote this approach on the blackboard during interactions with the students. Thus, the tasks and activities in the 'Pre-activities' focused on the strategies of calculation providing a motivation.

Phase IV (Exploration of proof ideas). After the teacher introduced a statement 'the sum of three consecutive numbers is divisible by three' in Phase III, students worked on three prototypical justifications referring to reasons why the statement is true. All justifications were expressed as verbal, figurative, or numerical representations rather than algebraic ones. For example, one justification given by a fictive character (Student 1) on the worksheet was as follows: ‘The sum is three times the number in the middle. Then I can divide this by three, without any remainder’. In this class, the teacher asked a student (Mila) to explain this idea.

T (00:43:59): Right, so [...] the sum is three times the number in the middle. Student 1 already showed this a moment ago. [...] Then I can divide this by three without a remainder. [...] Why does this work? […] Mila.

Mila (00:44:19): [...] Because you got, for example you got 8, and then on one side, you have got 7 and on the other side you have got 9. And for one you take something away and for the other one you put it on top. And then you can divide it by three, because then there are three equal numbers. [T writes “7+8+9” on the board, and underneath “8+8+8” and underneath “3*8”]

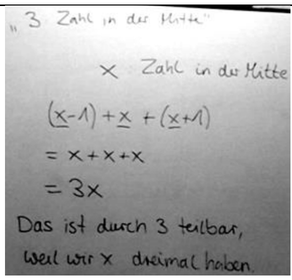
The teacher and Mila discussed the reasons underlying a general pattern such as ‘there are three equal numbers’ with examples (7+8+9=8+8+8=3*8). We could interpret these activities as the “exploration of proof ideas”.

Phase V (Formulation of a proof). Task 4 offers two different proofs (for G level). At the end of Phase V, the teacher wrote an additional proof with variables on the blackboard (Figure 2). The teacher constructed this proof by interacting with students and often asked them about the meaning or reason for the proof text. For example, the following is the excerpt regarding the meaning of “3x”. The teacher emphasized that it is important to write a text in order to understand the reasons behind it.

T (01:12:20): Right, [...] three times x. And now, someone has to explain to me, why this is divisible by three? [...] Divisible by three. Rodger.

Rodger (01:12:34): Because we have three times the same number.

T (01:12:36): Right. So, I have this on the board now, as a shortcut, I told many of you, you need to comment this, because like this, without what I have said, you do not really understand this. This is why I wrote it all down for you.



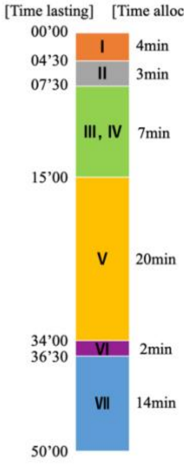
“3*middle number”
x   middle number

(x-1)+x+(x+1)
=x+x+x
=3x
This is divisible by 3,
because it is three times x.

Figure 2: A written proof on the blackboard (original in German, on the left)

Japanese lesson structure

Figure 3 presents an overview of a Japanese lesson structure comprising seven phases. There is a main task at the beginning of the lesson that guides the activities over the different phases. There are some key questions to describe the corresponding phases (see a table in Figure 3). We will illustrate some activities from phases I to VI.



Phases	Task/question descriptions	Categories
I	Task: For sum of three consecutive integers such as 1, 2, 3 or 6, 7, 8, conjecture a property that always holds.	Pre-activities
II	Question: Are there any properties?	Exploration and formulation of a statement
III	Question: How should we do for "it always holds"?	Exploration of proof ideas
IV	(No task/question)	
V	(Omitted – several small questions about the use of letters)	Formulation of a proof
VI	(No task/question)	Reflection
VII	(Working on a new statement: the sum of five consecutive numbers is divisible by five. And some additional tasks.)	Application activities

Figure 3: An overview of a Japanese lesson structure

Phases I (Pre-activities). In the first phase, the teacher clarified the meaning of 'three consecutive numbers' and then asked students to work on the task with examples. The students presented six examples: 0+1+2=3, 10+11+12=33, 100+101+102=303, etc. Therefore, 'pre-activities' are viewed as a moment to understand the given task.

Phases II (Exploration and formulation of a statement). In response to the teacher's question 'Are there any properties?', the students answered 'multiples of three'. The students found this inductively based on the examples. Then the teacher wrote it as a 'conjecture' on the blackboard. Subsequently, the teacher emphasised if the conjecture is always true, as the task refers to 'it always holds'.

T (05:02): The important part is that it "always holds." Does it always hold?

Nagi (05:04): Probably...

T (05:05): "Probably" is not enough. It is important that "it always holds." I want you to explain that "it always holds." What we have now is only six triads. […] Do you know how many numbers there are?

Hana (05:56): Infinite.

T (05:57): Yes, numbers exist infinitely many. We must explain about infinite numbers, because of "always."

As different activities related to the statement (such as discovering, formulating, and understanding the statement) were developed and interwoven, we identified two categories in this phase: 'exploration of a statement' and 'formulation of a statement'.

Phases III and IV (Exploration of proof ideas). The teacher asked 'How should we do (explain) that "it always holds"?'. Although the students explored this question in a group, nobody presented a possible way to explain. In Phase IV, the teacher suggested using letters and wrote on the blackboard 'To explain. Use letters!!'. This phase is also the moment to set up a learning goal of this lesson: 'Conjecture properties of numbers and explain that it always holds by using letters'. The conjecture (the sum of three consecutive integers is a multiple of three) has been already made in Phase II, so the teacher and the students came to work on the latter part of the goal. Thus, ‘proof ideas’ explored by them refer to using letters to prove the conjecture.

Phase V (Formulation of a proof). Figure 4 shows a proof written by the teacher on the board. The proof is called ‘explanation’ due to the curricular constraint; it counts as a standard proof that is commonly taught in Japanese classrooms. The proof was mainly led by the teacher, but some parts (e.g., how to express three consecutive numbers by using letters) were derived by interacting with the students. The teacher emphasised how the proof was structured. We could interpret, from the teacher’s remark, that the proof consisted of three parts: introduction, main body, and conclusion.

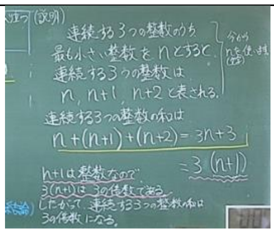
	(Explanation) Let n be the smallest integer of three consecutive integers, the three consecutive integers are represented as n, n+1, n+2. } From now on, we use n (letter) The sum of three consecutive integers is: n+(n+1)+(n+2)= 3n+3= 3 (n+1) Since n+1 is an integer, 3(n+1) is a multiple of three. (Conclusion) Therefore, the sum of three consecutive integers becomes a multiple of three.

Figure 4: A written proof on the blackboard (original in Japanese, on the left)

Phase VI (Reflection). The teacher mentioned that how to organise a proof was similar to how to organise a presentation in general: introduction, contents, and conclusion. At the end of this phase, the teacher concluded: ‘It’s easy to understand if you explain in this way. I think it always flows like this when we explain it. So try to look at the text in that way. This is also applicable in mathematics. OK? [36:00]’. We considered this as a meta-level comment on the proof which is included in ‘reflection.

DISCUSSION AND CONCLUSION

The characterizations of the lesson structures by the categories allowed us to better understand how proof and proving were organized and developed in the two lessons. This section presents and discusses some results from a comparative perspective. According to Figures 1 and 3, some common categories are organizing both lessons (pre-activities, formulation of a statement, exploration of proof ideas, formulation of a proof, and application activities), while our analysis shows some specificities of proof and proving in terms of the transitions between the categories. For example, the ‘pre-activities’ in the German lesson include the exploration of calculation strategies, that are connected to the activities in the ‘exploration of proof ideas’, in which the role of generic examples (Balacheff, 1987) is crucial. The ‘proof ideas’ in the German case include different representations for the justifications, which provide the reasons why the statement is true. In the Japanese case, some concrete examples, presented in the ‘pre-activities’, worked for discovering the conjecture, but not as generic examples to bring out the reasons why. The ‘proof ideas’ in this class only refer to the necessity of using letters. The emphasis is on how to prove that the given statement is *always* true. For the German case, proving are also the activities containing different approaches to justifications in the ‘formulation of a proof’ phase, which are developed from the ‘proof ideas’. Although the teacher only wrote the proof with variables on the board, another type of proof ('verbal proof') was included in Task 4. Regarding the written proof, the teacher explained that to write a text was to understand the reason behind it. For the Japanese case, the written proof (Figure 4) is the acceptable product in this class. Proving without the basis of generic examples is directly related to the construction of proof with variables. The written proof showed how a proof should be structured. This is also emphasized in the 'reflection' which is a category only found in the Japanese case. Regarding the 'application activities', both lessons involve an additional task applying the proof to an analogous statement (5 consecutive integers).

Our case study demonstrates how the seven categories, adapted from Boero (1999), can be used for a comparative approach to investigate proof lessons in different countries. This may contribute to advancing research on proof and proving from an international perspective (Reid et al., 2022). Based on the preliminary results, a deeper comparative analysis with theoretical and methodological elaborations is needed to understand which aspects are considered as cultural specificities in the two lessons.

Acknowledgement

This work was supported by JSPS KAKENHI (Grant Number JP20KK0053).

References

Balacheff, N. (1987). Processus de preuves et situations de validation. *Educational Studies in Mathematics*, *18*(2), 147–176.

Boero, P. (1999). Argumentation and mathematical proof: A complex, productive, unavoidable relationship in mathematics and mathematics education. *The Proof Newsletter*.

Brosseasu, G. (1997). *Theory of didactical situations, Didactique des mathematiqus 1970–1990*. Kluwer Academic Publishers.

Mariotti, M. A., Bartolini, M., Boero, P., Ferri, F., & Garuti, R. (1997). Approaching geometry theorems in contexts: From history and epistemology to cognition. In Pehkonen, E. (Ed.), *Proc. of 21st Conf. of the Psychology of Mathematics Education* (Vol. 1, pp. 180–195). PME.

Mariotti, M. A., Durand-Guerrier, V., & Stylianides, G. J. (2018). Argumentation and proof. In T. Dreyfus et al. (Eds.), *Developing research in mathematics education: Twenty years of communication, cooperation, and collaboration in Europe* (pp. 75–89). Routledge.

Reid, D., Shinno, Y, & Fujita, T. (2022). International perspective on proof and proving: Recent results and future directions. In C. Fernández. et al. (Eds.), *Proc. 45th Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 1, pp. 211–212). PME.

Stigler, J. W. & Hiebert, J. (1999). *The teaching gap: Best ideas from the worlds' teachers for improving education in the classroom*. Free Press.

Stylianides, G.J., Stylianides, A.J. & Moutsios-Rentzos, A. (2023). Proof and proving in school and university mathematics education research: a systematic review. *ZDM Mathematics Education*. https://doi.org/10.1007/s11858-023-01518-y

MAPPING COGNITIVE ENAGEMENT AND MOTIVATION: FINDINGS FROM THE ORRSEM PROJECT

Karen Skilling

University of Oxford, United Kingdom

The ORRSEM Project is concerned with secondary mathematics teachers' Observations, Recordings and Reports of Student Engagement and Motivation. A framework is presented that maps important motivational theories to types and levels of engagement, bringing achievement motivation and mathematics education research together. The findings from 4 teacher workshops sought teachers' descriptions of 41 engagement characteristics. Specifically, the 15 cognitive engagement characteristics are detailed because they are fundamentally valuable for educational outcomes, yet they are the least clearly conceptualized aspects of engagement research. The findings revealed that experienced teachers' are adept at identifying and describing the nuanced phases of self-regulation strategies and metacognitive processes.

INTRODUCTION

The influence of motivation and affective factors on student engagement for learning mathematics is considered as crucial and complex (Eccles, 2016). Although behavioural and overtly emotional engagement are more readily observed by teachers, subtle emotions and cognitive engagement are harder to identify and more difficult to clearly describe (Skilling et al, 2016). Cognitive engagement characteristics are focused on in this paper because they are the least clearly conceptualized aspect of engagement research (Skilling et al., 2016) despite being identified as fundamental to valuable educational outcomes (Michou et al., 2021) and performance (Lingel et al., 2019). One way to understand teacher engagement and motivation beliefs is through a research based and teacher informed tool, which provides a mechanism for noticing, articulating and communicating all types and levels of observed student engagement in mathematics classrooms. Understanding teachers' beliefs about how their students are engaging is relevant and important, as these shape instructional choices that can directly act to promote the extent to which students engage in mathematics learning and influence student outcomes, both of which are of concern in England.

The recently published Programme for International Student Assessments (PISA) 2022 Report for England results (Ingram et al, 2023), provide important details about the performance and mathematics experiences of English students, and have implications for teachers and teaching. In addition to capturing performance data, questionnaires sought students' views about their attitudes and beliefs and experiences in school, which acknowledges the role of affective and motivational factors that influence student engagement in learning settings. Compared to the OECD average of 75%, only 63% of English students felt that they belonged in school. The data also revealed 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. 4, pp. 105–112). PME.

36% of English students believed that their intelligence cannot be changed and regardless of how much study was done, they would not be good in mathematics or English. Although a healthy number of students (96%) want to do well in mathematics, only 44% of students reported mathematics was their favourite subject. The results emphasise the idiosyncratic nature of the mathematics learning experiences of many secondary students. Discrepancies and tensions are found, between wanting to do well yet not spending enough time on homework practice: or believing that mathematics is important, yet having a fixed mindset of their intelligence (Dweck, 2016).

What is it that teachers *can do* to more effectively support students' engagement and experiences with mathematics leaning? While the majority of students in England reported that their mathematics teachers provided extra help in mathematics lessons (80%), continued teaching until understanding is reached (72%), and showed interest in pupil learning (71%), it would be helpful to understand how teachers' know when and in what ways they can support individual students with their mathematics learning. This is critical for both improving learning outcomes and providing more positive mathematics experiences.

In many countries mathematics education is a core curriculum subject and is seen as a gateway subject for higher education study, STEM related careers and employment. Therefore, because engagement is crucial for encouraging student participation, interest, and learning in mathematics, it is important to investigate the extent to which teachers can identify all types of engagement. One way to elicit teacher' understanding of engagement is through a research based and teacher informed tool, which provides a mechanism for noticing, articulating and communicating all types and levels of engagement in mathematics classrooms. One of the aims of the ORRSEM Project is to refine such a tool that connects engagement characteristics more obviously and clearly link relevant motivational theories to specific types of engagement.

THEORETICAL BACKGROUND

In learning contexts, various sources of motivation are recognized as being reflected in student engagement such as: goals, needs, self-efficacy, control, value, and related responses such as fear of failure, avoidance, anxiety, mastery approaches and self-regulation. These factors are represented by various motivational theories and in combination, can indicate adaptive (positive) and maladaptive (negative) drivers of student engagement. The Engagement framework proposed by Fredricks, Paris and Blumenfeld (2004), which delineates behavioural (participation), emotional/affective (feelings, values, attitudes and interest) and cognitive (self-regulation and metacognition) types of engagement, is underscored by major motivational theories. Figure 1, represents a clear mapping of relevant motivation theories to types and levels of engagement and underpins the ORRSEM Teacher Tool.

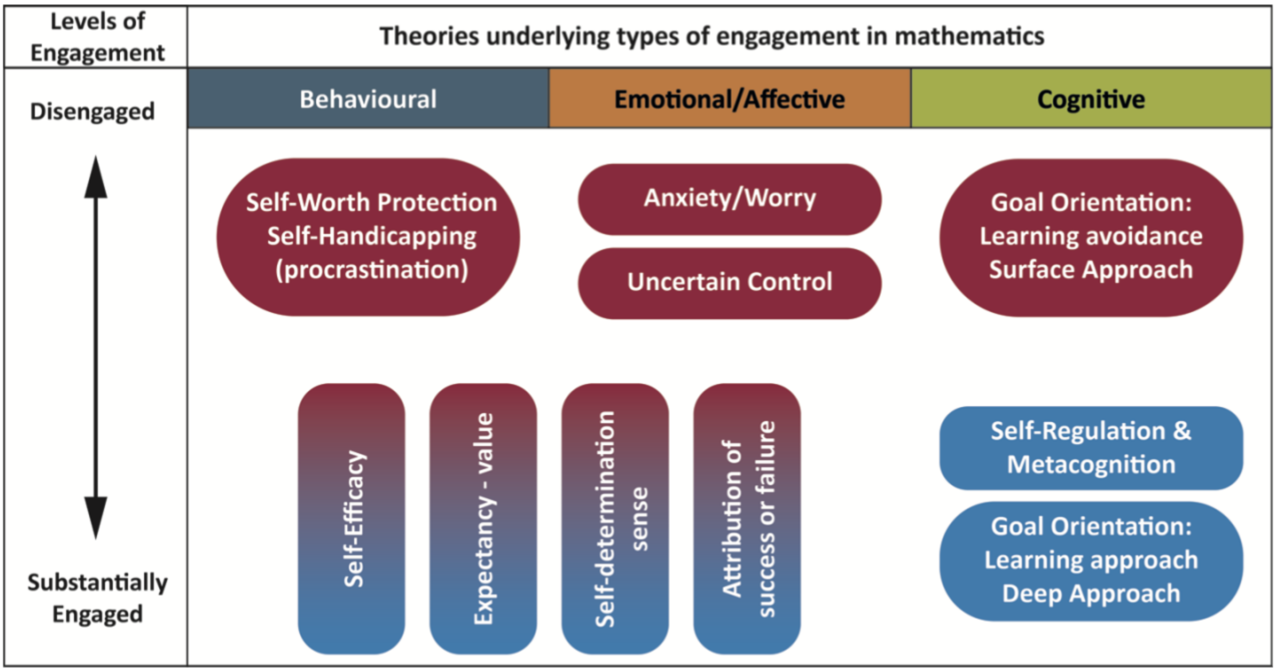


Figure 1: Mapping motivational theories to engagement

Looking specifically at *cognitive* engagement in the right hand side of Figure 1, this is mainly associated with the work of Pintrich (2004) and Zimmerman (2008), and includes self-regulation strategies for learning and metacognitive processes to describe the three main phases (forethought, performance, and reflection) and their sub-phases. The *forethought* phase refers to the activation of prior knowledge before commencing work such as setting goals; assessing what is known; and considering strategy use. Affective aspects are also relevant such as the valuing and interest in learning, and a student's self-efficacy for successfully reaching the goals set and what is driving their goals (intrinsic or extrinsic factors). The *performance phase* involves using and managing the strategies to reach desired goals (Pintrich 2004). Importantly, this includes efforts to modify, or change actions to maintain progress towards the goal and reduce possible distractions from reaching goals. This phase also includes monitoring progress through self-observation and keeping track of performance (Zimmerman, 2008). The *reflection phase* includes the self-reflection and learners' efforts to review and make judgements about their overall performance through feedback experiences with the task, including self-evaluation of what is being learned. Another aspect of this phase is the students awareness of how they are reacting in the learning process (their meta-awareness) and the effectiveness of their cognitive, affective, and behavioural choices (Zimmerman, 2008). While some students may spontaneously use a range of these subphases to manage their learning, many will not, therefore teachers can support students to become more skilful at autonomously self-regulating. The extent to which teachers' promote student engagement and autonomy for meeting students' innate psychological needs has been discussed by Reeve, (2009. Although autonomy is a separate construct from cognitive engagement, it is closely connected to motivational factors relevant to learning such as feeling a sense of belonging and competence.

In a study which focused on the cognitive engagement beliefs of secondary mathematics teachers, there was an almost equal split between teachers believing that their role was to *promote students to support their own learning* and those believing that they should *act as the support person for students* (Skilling & Stylianides, 2023). The main reasons given by teachers for their respective beliefs were connected to autonomy and its antithesis (control), for example: providing more or less structure/control to the students; and giving students more direction or information, versus encouraging student independence/autonomy. A Cognitive Engagement Framework (CEF) resulted from this study (Skilling & Stylianides, 2023), which conceptualized cognitive engagement through descriptions of phases and sub-phases relevant to self-regulation strategies and metacognitive processes. These phases and sub-phases are reflected in cognitive engagement characteristics (cells C1 to C15 in Figure 2) of the ORRSEM Teacher Tool where they are used in a mathematics specific context. The overall aims of this phase of the ORRSEM Project were to inform teachers about the different types of student engagement and associated motivation constructs, and refine the ORRSEM Teacher Template. The key research question asked: What engagement characteristics do mathematics teachers identify as important?

METHODOLOGY

Over an 18-month period, The ORRSEM Project worked with 16 teacher participants (T01-T16) from nine secondary schools (A-I) to better understand teacher' beliefs about, characteristics of, and practices for student engagement in mathematics. This paper reports details about the teacher participants descriptions of engagement characteristics which were mainly informed by the online workshops with small groups of teachers (the second phase of the research project). Details of the 14 of the teachers from eight of the schools (all state based schools in counties near to or in London) who took part in one of four workshops are recorded in Table 1.

Workshop #	Number of Participants	Schools (A-I)
# 1	5 (T: 01, 03, 07, 08, 15)	A, C, D, H
# 2	4 (T: 04, 12, 13, 14)	C, F, G, H
# 3	3 (T: 05, 06, 09)	C, D, E
# 4	2 (T: 02, 11)	B, F

Table 1: Workshop number, participants, and schools

The participants' years of teaching experience spanned between 2 and 20 years. Two teachers had 1-5 years' experience, three had 6-10 years' experience, and nine had over 10 years' experience. The participants attended an online workshop on a day that was convenient to them, resulting in four online workshops over a two week period with varying number of participants attending (see Table 1). The workshops lasted for 90 minutes. They began with a brief introduction by the researcher who outlined the definitions and conceptions of engagement in research literature and relevant underlying motivation theories. The participants were also introduced to a framework of engagement characteristics for each type of engagement: behavioural (n=13); emotional (n=13); and cognitive engagement (n=15). It was important that the engagement characteristics were considered in light of engagement theory (Fredricks et al, 2004) and underlying motivational theories. In this way it can be argued that construct validity (Harrits & Møller, 2021) is established.

The aim of the workshops was to use the characteristics from the prior research (Skilling et al, 2016) as a basis for the participant teachers to discuss their engagement beliefs and experiences in mathematics classrooms and reach a shared understanding of each characteristic. Importantly, the participants were invited to modify, refine, make additions or redactions of any characteristics they collectively agreed upon. In this way, each workshop group explicitly discussed engagement characteristics and reached an agreement on how each characteristic could be described and exemplified in ways that they deemed effective for mathematics teachers. Using the 'Board' function on Teams, each workshop group could annotate specific characteristics on the template provided. At the end of each workshop the annotated templates were downloaded and the recorded workshop conversations were transcribed for analysis.

Approach to analysis

The researcher scrutinized the contributions each workshop group made for each of the characteristics. Print outs of the workshop boards were compared to note similarities and differences in descriptions. Then the researcher distilled the contributions of all four workshop groups to summarise the teachers' reports for each of the 41 engagement descriptors. A second researcher was asked to interpret the summaries made by the lead researchers and randomly chose 10% of the summaries for checking. They reached a 91.67% agreement which is within the recommended rater-reliability (Krippendorf, 2004). The next section reports the distilled responses by each type of engagement and provides justification for the resulting ORRSEM framework.

FINDINGS

The number of responses for each of the 41 descriptors for each type of engagement were similar. Specifically, for *behavioural* engagement, there were an average of 33 responses per workshop; for *emotional/affective* the average number of responses was 31.3; and for *cognitive* engagement the average was 33. As mentioned, the findings in this paper will focus on the *cognitive* engagement descriptions from the teacher workshops which are presented in Figure 2. A total of 100 descriptions were made for *cognitive* engagement and were evenly spread between cells C1 to 15.

Cognitive Engagement Descriptors	Summarised Teacher Responses
C1 Activating knowledge - Finds it hard to recall	**Forgets** facts; **fails to recall** prior material despite prompts; **guessing**.
C2 Strategy choice - Simplistic	Tries every problem **the same way**; does **not use strategies** provided or follow direct instructions or; looks at the work of others.
C3 Strategy planning - Not sure when and why to apply knowledge	Uses formula in the same way **instead of finding more efficient methods**; has **difficulties making connections** between concepts; has difficulties in spotting patters and similarities; **not flexible** when responding to 'applied' questions; requires **constant prompting**.
C4 Regulation of thinking - Rarely evident	Completes tasks **without understanding**; **not knowing which path to take**; reflected by who stays on task and **who gives up**; requires **telling students exactly what mathematics to use** to answer the question.
C5 Monitoring learning - Rarely evident	Not realising how much **time is spent on one question**; **robotic approach** to self-checking or self-marking; does **not use previous work**; **repeats previous mistakes**; does **not complete homework**.
C6 Refection on learning – rarely evident	Self-marks with **no care about solutions**; copies material from board for the 'right' solutions; **forgets previous learning** and **unable to apply it**; complete questions but do **not mark or check** them; does **not engage with feedback**.
C7 Achievement orientation – mainly on performance	**Focus on getting questions correct rather than the method/strategy;** compare themselves with others in the class; very competitive; stress about assessments; is discouraged with low marks.
C8 Does enough to avoid failure	Do **not challenge themselves**; **chooses easy options**, not extension tasks; asks how much/many questions have to do; **does the minimum/sufficient**.
C9 Activating knowledge – fluently recalls	Engages in starter questions; asks questions; uses a **wide range of strategies**; understands the **steps to find solutions**; quick to **link topics** and **uses methods** for this; **able to recall formula** and can work without a calculator.
C10 Strategy choice – usually appropriate	Looks for the best method and **approach for an effective solution**; can provide **reasons for choosing particular strategies**; can **justify answers**; can **identify correct route from multiple strategies.**
C11 Strategy planning– knows when and why to apply knowledge	Can **answer questions in any context**; can **utilise knowledge across different mathematics strands**; can **identify best approach to answering** questions; **uses multiple strategies to solve**; is **able to help and guide other** students.
C12 Regulation of thinking– often evident	Talking with **other students and debating work**; ask **questions ahead of class**; **refers to books**; **discusses reasoning**; asks questions that reflect **flexible thinking**; **talk through workings**; **check answers**; have **an end goal**.
C13 Monitoring learning– often evident	Takes care with marking and **whether right or wrong; identifies learning progress**; **progress measured through understanding** not task completion; chooses to try **different ways of approaching questions**; keeping corrections and **checking for improvement** over time; engaging in **feedback and self-assessment**; looks to **identify areas of improvement**; how current **work fits with other areas of maths.**
C14 Refection on learning – often evident	**Cares about what is wrong** and makes corrections; **not repeating same mistakes**; corrects work and **learns different and effective methods**; **asks how to improve**; **moments of realisation.**
C15 Achievement orientation – mainly for mastery	Wanting to **understand why** and **reason through the solutions and methods that work**; **seek out problems to solve;** seeks the **boundaries of their knowledge;** not being content to regurgitate a method without understanding.

Figure 2: Cognitive Engagement Descriptions - Summarised teacher responses

It can be seen from Figure 2 that different shades have been used to indicate broad levels of engagement. Cells C1 to C6 indicate low engagement or disengagement; cells C7 to C10 indicate varying or non-substantial engagement; cells C11 to C15 indicate substantial levels of engagement. In the 'Summarised Teacher Responses' section of Figure 2, particular words and phrases have been bolded. These reflect the phases and subphases of self-regulation and metacognition discussed in the relevant literature.

As expected there are more positive mentions of specific self-regulation and metacognitive activities at the substantial engaged range (C9 to C15) of Figure 2, however it is particularly valuable that the teachers' identified characteristics that were missing from, or acted to, undermine students' cognitive engagement in cells C1 to C8.

For example, dominant characteristics between C1 and C8 included not activating knowledge ('failing to recall', 'not knowing what path to take', 'guessing', 'looking at the work of others', 'difficulties making connections'). Aspects relevant to performance were also identified such as, not enacting strategies or regulating (e.g. 'using the same methods', 'trying every problem in the same way', 'spending time on single questions'). In terms of reflection, a 'robotic approach to self-checking', lack of care about the solutions, 'repeating the same mistakes' and 'not engaging in feedback' was reported. Reports relevant to goal orientation indicates this was performance driven, focusing on correct questions rather than mastering the method or strategy.

Some self-reaction comments in the variably engaged cells, reflected affective aspects such comparing oneself to classmates, being competitive, ‘stressed about assessments’ and ‘discouraged with low marks’.

In contrast, the main characteristics reported between cells C9 and C15 included substantial activation of knowledge and ‘making links between topics’ and ‘across different mathematics strands’, as well as ‘recalling formula’. Having an ‘end goal and not being distracted from it’ was recounted. Numerous mentions of strategy use (choosing between wide and multiple strategies), for working through solutions and understanding methods. This was evidenced through ‘talking with other students and debating work’, ‘asking questions ahead of class’, ‘discussing reasoning’. Goal orientation was achievement focused, indicated by ‘caring about whether right or wrong’, understanding the steps and best approach for solutions and ‘seeking the boundaries of knowledge’. Reflection characteristics were also evident such as ‘engaging in feedback and self-assessment’, ‘not repeating mistakes’ and ‘moments of realisation’. Typically, characteristics of cognitively engaged students, can quickly identify and recall relevant knowledge, knowing when and why to use particular strategies, checking their progress on tasks and overall improvement, responding to feedback, asking questions and reflect on their learning goals and their understandings.

DISCUSSION AND CONCLUSION

Although teacher practitioners may not be knowledgeable about motivation theory and constructs they recognise the effects of motivational and affective factors by the patterns of behaviours and reactions observed in mathematics classrooms (Skilling, 2016). It is crucial to better inform teachers about the possible motivational factors driving student engagement for several reasons. Teachers may become more alert to signals that indicate shifts in a student’s behavioural, emotional and cognitive attention. It may also assist teachers to identify the diversity of individual students’ engagement patterns, resulting in teachers’ being better equipped to shape student support.

In the results reported in this paper, it is evident that the teachers were adept at identifying types of cognitive actions and engagement and articulating these by drawing on their experiences in mathematics classrooms. The nuances descriptions by the teachers resonates with every phase and subphase that are described in self-regulation and metacognition literature, and included affective factors that are associated with psychological investment in learning and student efforts. These findings are in contrast to previous studies where the reduced scope and frequency of cognitive characteristics indicated this aspect of engagement was not well understood by teachers (Skilling et al., 2016). A key difference in the ORRSEM study was the planned phases of the study to include collaborative teacher workshops to draw on teacher knowledge and experiences. The opportunity to discuss the association between engagement and motivation factors before the workshops began may well have assisted with teacher clarity. The opportunity to collaborative discuss student engagement with other teachers and identify aspects specific to mathematics was crucial and resulted in mathematics specific engagement characteristics and an important contribution to mathematics education research. An obvious limitation is the scale of the project, however a specification for an ORRSEM app for digital recoding of teachers engagement observation has been prepared and funding is being sought.

Additional information

The ORRSEM Project received John Fell Funding. I thank Sara Berkai and Cindy Wong who acted as Research Assistants for different phases of the project.

References

Dweck, C. (2016), Mindset: The New Psychology of Success, Ballantine Books, New York.

Eccles, J.S. (2016). Engagement: where to next? Learning and Instruction, 43, 71–75.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H., (2004). ‘School engagement: potential of the concept, state of the evidence’, *Review of Educational Research* 74, pp. 59-109, 2004.

Harrits, G.S., & Møller, M. Ø., (2021). Qualitative vignette experiments: a mixed methods design, *Journal of Mixed Methods Research,* 15 f(4), 5 26-545, 2021.

Ingram, J., Stiff, J., Cadwallader, S., Lee, G., & Kayton, H. (2023). PISA 2022: National Report for England.

Krippendorf, K. (2004). *Content analysis: An introduction to its methodology* (2nd ed). Thousand Oaks, CA: Sage.

Lingel, K., Lenhart, J., & Schneider, W. (2019) Metacognition in mathematics: Do different metacognitive monitoring measures make a difference? *ZDM – Mathematics Education, 51*, 587–600.

Michou, A., Altan, S., Mouratidis, A., Reeve, J.M., & Malmberg, L.E., (2021). Week-to-week interplay between teachers’ motivating style and students’ engagement, *The Journal of Experimental Education*.

Pintrich, P. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Education Psychology Review*, *16*, 385–407.

Reeve, J. (2009). Why teachers adopt a controlling motivating style toward students and how they can become more autonomy supportive. *Educational Psychologist*, *44*(3), 159–175.

Skilling, K., Bobis, J., Martin, A. J., Anderson, J., & Way, J. 2016, ‘What secondary teachers think and do about student engagement in mathematics’, *Mathematics Education Research Journal 28*(4), pp. 545-566.

Skilling, K., & Stylianides, G. J. (2023). Using vignettes to investigate mathematics teachers’ beliefs for promoting cognitive engagement in secondary mathematics classroom practice. ZDM – The International Journal on Mathematics Education, 55, 477-490

Zimmerman, B. J. (2008). Investigation self-regulation and motivation: Historical background, methodological development, and future prospects. *American Educational Research Journal*, *45*, 166–183.

AGE MATTERS WHEN IT COMES TO STUDENTS' ATTITUDES TOWARD ONLINE MATHEMATICS ASSESSMENTS

Erica Dorethea Spangenberg

University of Johannesburg

This study established how students' attitudes toward online mathematics assessments relate to age due to the shift to online learning during COVID-19. Quantitative data were collected through an adapted Attitudes Toward Mathematics Inventory from 734 students in seven South African schools. Although enjoyment, perceived usefulness, ease of use, and self-confidence in engaging in online mathematics assessments decline with age, they are significantly lower for students 13-16 years old compared to those older (17-22 years) and younger (10- 12 years). Intrinsic motivation is statistically the same for older and younger students but significantly lower for students who are 13-16 years old. This study suggests further research on affective aspects influencing specific types of online mathematics assessments.

INTRODUCTION

During COVID-19, several studies have been conducted on the affordances and challenges of online teaching in mathematics (Engelbrecht et al., 2020). However, Collimore et al. (2015) suggested further research on students' attitudes toward online assessments to ensure students' engagement in and learning from online assessment tools. Focusing on students' attitudes toward online assessments is essential, as positive attitudes toward online assessments may improve students' enjoyment, self-confidence, and motivation when engaging with mathematics content, which may determine how to proceed with online assessments. Besides, the age of learners may also play a role in attitudes, especially considering Zuo et al. (2021), who found that age significantly affects perceived online learning experiences. Consequently, the study aimed to establish how students' attitudes toward online mathematics assessments relate to age.

THEORETICAL FRAMEWORK

The self-determination theory coined by Deci and Ryan (1985) theoretically underpinned the study. This theory suggests that students have three basic psychological urges (competence, autonomy, and relatedness) to fulfill to ensure effective functioning and wellness (Deci & Ryan, 2015). Within this theory, (1) enjoyment, as a symptom of wellness, is viewed as an essential positive attitude related to intrinsic motivation; (2) intrinsic values are associated with wellness; and (3) social and environmental factors satisfying psychological inclinations for autonomy, competence, and relatedness, associated with self-confidence and perceived usefulness, expedite enjoyment. For this study, attitudes are inclinations and aversions toward online assessments in mathematics consisting of five dimensions: (1)

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. 4, pp. 113–120). PME.

enjoyment; (2) self-confidence; (3) perceived usefulness; (4) ease of use; and (5) intrinsic motivation, which structure the conceptual framework. The attitudes toward mathematics inventory (ATMI) (Tapia & Marsh, 2004) also foregrounded these attitudinal dimensions, which were adapted to gather data for this study.

Students are driven by emotions when they experience challenging undertakings like being assessed online in mathematics. According to Yilmaz et al. (2022), if students enjoy such tasks, they cope, stay interested, and perform in mathematics and are intrinsically motivated to develop goal-orientated behaviour toward mathematics. Current studies focusing on enjoyment in online mathematics assessment at the school level are scant despite several studies focusing on enjoyment in mathematics learning (Rodríguez et al., 2021). Yilmaz et al. (2022) also acknowledged that online learning is barely a flavoured undertaking at the school level, and research on emotions in online learning is limited. A study by Rodríguez et al. (2020), which could also apply to online assessments, found that students from 13 primary schools in Spain who were 9-13 years old experienced mathematics learning as pleasant and perceived mathematics content as valuable. The participants also believed they were competent in doing mathematics.

Self-confidence in online assessments in mathematics refers to a person's assurance of being competent to excel in mathematics online tests compared to others (Bringula et al., 2021). In addition, self-confidence in using online tools ensures performing well in online mathematics assessments (Acosta-Gonzaga & Walet, 2018). Thus, students' self-confidence to engage in mathematics activities and the ease of using online tools may impact their success in mathematics being assessed online. However, students' self-confidence in being assessed online in mathematics may change as they progress in grades. This sentiment aligns with Foster et al. (2022), who conducted diagnostic multiple-choice mathematics questions online with 7302 UK primary and secondary school students and found that students' self-confidence declines as they age.

Tapia and Marsh (2004) identified utility value referring to the significance of an endeavour or product to be useful or prominent in acquiring a preferred result, as an attitudinal dimension towards mathematics. However, Zuo et al. (2021) separated this dimension further into two different constructs – perceived usefulness and ease of use, which were also adopted for this study. Perceived usefulness signifies beliefs that online learning enhances performance, while ease of use implies that using technological tools to learn is effortless. Therefore, instead of signifying utility value, this study foregrounded perceived usefulness as students' views about online assessments providing evidence of their competence in mathematics and ease of use as students' perceptions of the easiness or difficulty they encountered when engaging in online assessments in mathematics. Zuo et al. (2021) found that the perceived usefulness of online teaching mainly contributed, while the ease of technological tools played a moderate role in the overall learning engagement of 118.589 Chinese students. Age had the most significant impact on perceived online learning experiences, and school location reflected vast inequities. At the same time, there were no significant differences in terms of gender and the use of technology tools. Opposingly, Thurm et al. (2022) found that secondary school students from Flanders, Germany, and the Netherlands view the utility value of mathematics as moderate. They revealed that students who received synchronous teaching with high-order didactical approaches had a conducive environment to study and portrayed positive attitudes toward mathematics. These students reported more positive beliefs about online assessments than those from educationally disadvantaged backgrounds. Findings from these studies may suggest that students' perceived usefulness during online assessments decreases as they progress in age, especially in the light of Foster et al. (2022), who revealed a decline in students' attitudes toward mathematics during the administration of assessments in an online space in the transition from primary to secondary school.

Intrinsic motivation refers to undertakings executed for their purpose, deep-rooted significance, and pleasure (Liu, 2021). Scherrer and Preckel (2019) added that intrinsic motivation is not controlled or strengthened by sources outside a person, thus externally motivated, but through amusement, discovery, and own actions. As intrinsic motivation has proved to be an essential predictor for learning success in mathematics (Pelikan et al., 2021) despite a significant decline in its levels as the students age (Scherrer & Preckel, 2019), the dimension is subsumed as essential to establish students' attitudes toward online assessments in mathematics. Pelikan et al. (2021) found that students who perceived themselves as competent in mathematics during COVID-19 when they had to learn online, were more intrinsically motivated than those who believed they were less competent. On the other hand, Liu (2021) revealed that avoiding challenging activities in mathematics, such as online assessments and failure, may be detrimental to intrinsic motivation to engage in mathematics.

In summary, Rodríguez et al. (2021) found that students' perceived usefulness is related to their perceptions about their competence in mathematics and intrinsic motivation to engage in the subject, and, thus, the key to their self-confidence, enjoyment, and well-being when learning mathematics. Hence, the following null hypotheses were tested in the study:

- H_0: There is no significant difference in the enjoyment of online mathematics assessments of learners in different age groups.
- H_0: There is no significant difference in self-confidence towards online mathematics assessments of learners in different age groups.
- H_0: There is no significant difference in perceived usefulness towards online mathematics assessments of learners in different age groups.
- H_0: There is no significant difference in ease of use towards online mathematics assessments of learners in different age groups.
- H_0: There is no significant difference in intrinsic motivation towards online mathematics assessments of learners in different age groups.

METHOD

This exploratory study adopted a qualitative method. A survey comprising 40 Likert-type scale items with a five-point scale amended from the Attitudes Toward Mathematics Inventory (ATMI) for high school students (Tapia & Marsh, 2004) was distributed to participants. The instrument comprised five dimensions: enjoyment; self-confidence; perceived usefulness; ease of use; and intrinsic motivation. The ATMI was chosen for its proven evidence of content validity to measure attitudes toward mathematics in several research studies for different age groups and contexts (Afari, 2013; Anastasiadis & Zirinoglou, 2022; Simegn & Asfaw, 2017).

Seven hundred and sixty-two participants from upper primary (10-12 years old), lower secondary (13-16 years old), and upper secondary (17-22 years old) school levels in seven South African schools were randomly selected based on voluntary participation and who had experience of online mathematics assessments during COVID-19. Surveys were distributed to them for completion.

Twenty-eight students' collected data were omitted due to missing information. Thus, data from 734 questionnaires were analysed using the statistical software package SPSS version 28. Eleven negatively worded items were reverse coded to ensure consistent questions' directionality and prevent distorted responses from participants.

The ATMI is an established standarised instrument and conforms with construct and content validity to measure students' attitudes toward mathematics. However, to confirm face validity, eight teacher researchers peer-reviewed and modified the instrument to ensure that the participants would accurately understand the items regarding the online mathematics assessments and that the language was accessible within the study's context. Internal validity was addressed by employing Principal Components Analysis (PCA), and reliability was established by determining Cronbach's α coefficients.

RESULTS

PCA was performed in an exploratory manner to discover the shared attitudinal dimensions toward online mathematics assessments, as the wording of the items of the ATMI was changed to fit the study's context. The Kaiser–Meyer–Olkin (KMO) test (KMO value = .96) confirmed the sampling adequacy and showed a marvellous sample for the factor analysis. Bartlett's test of sphericity (χ^2 (666) = 13017, p = .0000 < .001) indicated that the correlations between items were adequately significant for PCA. The internal structure of the 40 residual items was tested with orthogonal rotation (varimax) with Kaiser normalisation. The rotation converged in seven iterations. Two factors loaded less than .40, and one item did not correlate with the other variables. Therefore, these three variables were eliminated. The highest loading was considered for items loaded simultaneously on multiple factors. The Kaiser–Guttman rule was used to identify the components with a related eigenvalue greater than one. Five attitudinal dimensions had eigenvalues greater than the Kaiser's criterion of 1 and, added together, accounted for 54.4% of the variance and were thus used for analysis. The items were assessed, and subsequently, clusters of items were named as they naturally transpired and according to the researcher's interpretation of the items, namely enjoyment (eight items), self-confidence (eight items), perceived usefulness (six items), ease of use (ten items), and intrinsic motivation (five items).

The Cronbach's alpha (α) coefficient for the 37 test items was .95. Thus, the internal consistency of the scale items was assumed to be excellent. Respectively, all five attitudinal dimensions had good reliability: enjoyment (α =.88), self-confidence (α = .86), perceived usefulness (α = .85), ease of use (α = .89), and intrinsic motivation (α = .83).

A Shapiro-Wilk test was employed to assess normality. As the data across the five attitudinal dimensions were non-parametric (significant value smaller than .05), a Kruskal-Wallis test was performed to establish whether statistically significant differences existed between the age groups. Attitudes toward online assessments in mathematics were significantly lower for the participants who were 13-16 years old compared to older (17-22 years) and younger (10-12 years) participants across four attitudinal dimensions:

- Enjoyment ($H(2, n = 734) = 17.5$, $p < .001$) with a mean rank of 426 (Md = 29.0) for participants 10-12 years old, 342 (Md = 25.0) for participants 13-16 years old, and 402 (Md = 28.0) for participants 17-22 years old.
- Perceived usefulness ($H(2, n = 734) = 25.4$, $p < .001$) with a mean rank of 467 (Md = 22.0) for participants 10-12 years old, 341 (Md = 19.0) for participants 13-16 years old, and 389 (Md = 20.0) for participants 17-22 years old.
- Ease of use ($H(2, n = 734) = 48.8$, $p < .001$) with a mean rank of 500 (Md = 39.0) for participants 10-12 years old, 329 (Md = 31.0) for participants 13-16 years old, and 405 (Md = 35.0) for participants 17-22 years old.
- Intrinsic motivation ($H(2, n = 734) = 26.4$, $p < .001$) with a mean rank of 433 (Md = 16.0) for participants 10-12 years old, 336 (Md = 15.0) for participants 13-16 years old, and 412 (Md = 16.0) for participants 17-22 years old.

For self-confidence, one non-empty group prevented the researcher from performing a Kruskal-Wallis test on that dimension and opting for a Mann-Whitney U test. The attitudes toward online assessments in mathematics for students who were 10 -12 years old regarding self-confidence were higher than those in the other two age groups, with participants who were 13-16 years old having the lowest self-confidence:

- participants 10-12 years old (Md = 31.0) versus participants 13-16 years old (Md = 24.0), U = 7285, $z = -7.230$, $p = .001 < .05$, $r = -.27$ (a finding with a small practical significance)
- participants 10-12 years old (Md = 31.0) versus participants 17-22 years old (Md = 25.0), U = 4021, $z = -5.859$, $p = .001 < .05$, $r = -.22$ (a finding with a small practical significance).

- participants of 13-16 years (*Md* = 24.0) versus participants 17-22 years old (*Md* = 25.0), $U = 43101$, $z = -2.236$, $p = .025 < .05$, $r = -.08$ (a finding with a small practical significance).

DISCUSSION

Although lower than the other age groups, participants 13-16 years had the highest median for ease of use, followed by enjoyment, perceived usefulness, and intrinsic motivation. This finding differs slightly from Thurm et al. (2022), who found that secondary school students perceive the utility value of online assessments in mathematics to be moderate. Zuo et al. (2021) also revealed that the perceived usefulness of online teaching mainly contributed, while the ease of using technological tools played a moderate role in the overall learning engagement of their sample. The students might have valued online assessments in mathematics easier as they might have practised assessments before the actual assessment and could change answers before the final submission for assessment.

Younger participants (10-12 years) enjoyed online mathematics assessments more, viewed themselves as better competent, and found online assessments easier to use than the older participants (17-22 years). This finding opposes Acosta-Gonzaga and Walet (2018), noting that students experiencing feedback from online assessments is more enjoyable than helpful. Young students may enjoy online assessments in mathematics if they are exposed to fun assessments, for example, mathematics games, to be completed. However, older students might have been less flexible in adapting to online assessments in mathematics, as they were comfortable with being assessed face-to-face like in the past.

The positive correlation between younger (10-12 years) and older participants (17-22 years) being intrinsically motivated aligns with Acosta-Gonzaga and Walet (2018), revealing feedback to be crucial in online mathematical assessments, as it affects performance in mathematics positively and stimulates students' intrinsic motivation. The students could be intrinsically motivated by online assessments in mathematics because they viewed the online assessment platform as non-judgemental and less intimidating and received prompt feedback informing them about their progress at an early stage. The students might have also felt more comfortable and in control by completing their mathematics assessment online at their own pace and place.

The higher self-confidence toward online assessments in mathematics of younger students compared to the other two age groups aligns with Foster et al. (2022), revealing that students' self-confidence declines as they progress with age. The lower self-confidence regarding online mathematics assessments of participants who were13-16 years old but higher self-confidence in younger and older students could be ascribed to either negative experiences when studying mathematics in the past or students being in the process of finding a balance between experience and expectations on being competent in mathematics, which may change once they have developed an identity.

CONCLUSION

Despite ample recent research on the advantages and shortcomings of online mathematics teaching and learning, studies on students' attitudes toward online mathematics assessments and how they relate to age are limited. Therefore, this study established how students' attitudes toward online mathematics assessments relate to age. Although enjoyment, perceived usefulness, ease of using, and self-confidence in engaging in online assessments in mathematics decline with age, they are significantly lower for students 13-16 years old compared to those older and younger. Intrinsic motivation is statistically the same for older and younger students but also significantly lower for students 13-16 years.

Extending this research to other contexts may result in more robust support for the findings. In-depth interviews and observations while engaging in online mathematics assessments can complement and strengthen the study's results. A longitudinal study focusing on changes in students' attitudes towards specific types of online mathematics assessments will also supplement the results.

This study contributes to current research on affective aspects influencing online mathematics assessment. Although it is complex to comprehend the manifestation of attitudes, this study offers a starting point for understanding how attitudinal dimensions relate to different age groups. It upholds the argument that students' attitudes toward online assessment in mathematics become negative as they become adolescents (13-16 years).

References

Acosta-Gonzaga, E., & Walet, N. R. (2018). The role of attitudinal factors in mathematical online assessments: a study of undergraduate STEM students. *Assessment & Evaluation in Higher Education*, *43*(5), 710–726. https://doi.org/10.1080/02602938.2017.1401976

Afari, E. (2013). Examining the factorial validity of the attitudes towards mathematics inventory (ATMI) in the United Arab Emirates: Confirmatory factor analysis. *International Review of Contemporary Learning Research*, *2*(1), 15–29. https://doi.org/10.12785/irclr/020102

Anastasiadis, L., & Zirinoglou, P. (2022). Students' attitudes toward mathematics: The case of Greek students. *International Journal of Education and Social Science*, *9*(3), 60–76.

Bringula, R., Reguyal, J. J., Tan, D. D., & Ulfa, S. (2021). Mathematics self-concept and challenges of learners in an online learning environment during COVID-19 pandemic. *Smart Learning Environments*, *8*(1), 1–22. https://doi.org/10.1186/s40561-021-00168-5

Collimore, L.-M., Paré, D. E., & Joordens, S. (2015). SWDYT: So what do you think? Canadian students' attitudes about peerScholar, an online peer-assessment tool. *Learning Environments Research*, *18*(1), 33–45. https://doi.org/10.1007/s10984-014-9170-1

Deci, E. L., & Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. *Journal of Research in Personality*, *19*(2), 109–134. https://doi.org/10.1016/0092-6566(85)90023-6

Deci, E. L., & Ryan, R. M. (2015). Self-determination theory. In *International Encyclopedia of the Social & Behavioral Sciences* (pp. 486–491). Elsevier. https://doi.org/10.1016/B978-0-08-097086-8.26036-4

Engelbrecht, J., Llinares, S., & Borba, M. C. (2020). Transformation of the mathematics classroom with the internet. *ZDM*, *52*(5), 825–841. https://doi.org/10.1007/s11858-020-01176-4

Foster, C., Woodhead, S., Barton, C., & Clark-Wilson, A. (2022). School students' confidence when answering diagnostic questions online. *Educational Studies in Mathematics*, *109*(3), 491–521. https://doi.org/10.1007/s10649-021-10084-7

Liu, W. C. (2021). Implicit theories of intelligence and achievement goals: A look at students' intrinsic motivation and achievement in mathematics. *Frontiers in Psychology*, *12*, 1–12. https://doi.org/10.3389/fpsyg.2021.593715

Pelikan, E. R., Lüftenegger, M., Holzer, J., Korlat, S., Spiel, C., & Schober, B. (2021). Learning during COVID-19: The role of self-regulated learning, motivation, and procrastination for perceived competence. *Zeitschrift Für Erziehungswissenschaft*, *24*(2), 393–418. https://doi.org/10.1007/s11618-021-01002-x

Rodríguez, S., Estévez, I., Piñeiro, I., Valle, A., Vieites, T., & Regueiro, B. (2021). Perceived competence and intrinsic motivation in mathematics: Exploring latent profiles. *Sustainability*, *13*(16), 8707. https://doi.org/10.3390/su13168707

Rodríguez, S., Regueiro, B., Piñeiro, I., Valle, A., Sánchez, B., Vieites, T., & Rodríguez-Llorente, C. (2020). Success in mathematics and academic wellbeing in primary-school students. *Sustainability*, *12*(9), 3796. https://doi.org/10.3390/su12093796

Scherrer, V., & Preckel, F. (2019). Development of motivational variables and self-esteem during the school career: A meta-analysis of longitudinal studies. *Review of Educational Research*, *89*(2), 211–258. https://doi.org/10.3102/0034654318819127

Simegn, E. M., & Asfaw, Z. G. (2017). Assessing the influence of attitude Towards mathematics on achievement of grade 10 and 12 female students in comparison with their male counterparts: Wolkite, Ethiopia. *International Journal of Secondary Education*, *5*(5), 56–69. https://doi.org/10.11648/j.ijsedu.20170505.11

Tapia, M., & Marsh, G. E. (2004). An instrument to measure mathematics attitudes. *Academic Exchange Quarterly*, *8*(2), 16–21.

Thurm, D., Vandervieren, E., Moons, F., Drijvers, P., Barzel, B., Klinger, M., van der Ree, H., & Doorman, M. (2022). Distance mathematics education in Flanders, Germany, and the Netherlands during the COVID-19 lockdown—the student perspective. *ZDM – Mathematics Education*. https://doi.org/10.1007/s11858-022-01409-8

Yilmaz, M. B., Orhan, F., & Zeren, S. G. (2022). Adolescent emotion scale for online lessons: A study from Turkey. *Education and Information Technologies*, *27*(3), 3403–3420. https://doi.org/10.1007/s10639-021-10734-6

Zuo, M., Ma, Y., Hu, Y., & Luo, H. (2021). K-12 students' online learning experiences during COVID-19: Lessons from China. *Frontiers of Education in China*, *16*(1), 1–30. https://doi.org/10.1007/s11516-021-0001-8

THE ROLE OF MATHEMATICS AND INSTRUCTIONAL PRACTICES IN INTEGRATED STEM EDUCATION

Carina Spreitzer, Verena Kaar, David Kollosche, and Konrad Krainer

University of Klagenfurt

The integration of science, technology, engineering, and mathematics (STEM) in education has gained momentum, driven by the acknowledgment that real-world challenges demand a holistic approach. This study explores the intersection of integrated STEM and mathematic. The research assesses eleven existing materials according to instructional practices and the role of mathematics. Results reveal a comprehensive incorporation of STEM instructional practices. However, the role of mathematics is often used as an ancillary discipline, employed primarily as a tool in STEM activities. Only a minority of materials explicitly integrate mathematical concepts within interdisciplinary contexts. The findings underscore the need for a more pronounced role of mathematics in integrated STEM education.

INTRODUCTION

Czerniak et al. (1999) already stated that curriculum integration in science and mathematics has become popular in the 1990s. They substantiate this hype with the argument that “in the real world, people’s lives are not separated into separate subjects; therefore, it seems logical that subject areas should not be separated in schools” (p. 421). Also recently published reviews underpin the increasing research interest in Science, Technology, Engineering, and Mathematics (STEM) education (e.g., Bozkurt et al., 2019; Kurniati et al., 2022; Le Thi Thu et al., 2021). The main objective of this research is to address global issues. The Education 2030 initiative underscores the increasing difficulties schools face in equipping students for rapid economic, environmental, and social transformations, as well as for emerging jobs and technologies, while also tackling unforeseen social issues (OECD, 2018). In response to these challenges and the need for an increasing number of STEM graduates, Bøe et al. (2011) contend that having qualified STEM professionals is crucial for maintaining economic competitiveness. An encouraging strategy in this regard is the adoption of an integrated STEM curriculum, providing learners with opportunities for “more relevant, less fragmented, and more stimulating” (Furner & Kumar, 2007, p. 186) experiences. Secondary school plays a crucial role in the STEM orientation of students (Reinhold et al., 2018) and especially the success in mathematics during secondary school is a further factor in fostering students for a STEM career (e.g., Kohen & Nitzan-Tamar, 2021; Nitzan-Tamar & Kohen, 2022). So therefore, the interplay between integrated STEM and mathematics is essential in answering the challenges in the next years.

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. 4, pp. 121–128). PME.

More and more schools, educational administrations and countries are striving for an interdisciplinary STEM subject. The "MINT-MS" (STEM middle school) school pilot project was introduced in Austria with the start of the 2022/2023 school year. As part of this pilot project, the secondary school curriculum, which in Austria covers grades 5-8, has been expanded to include an integrated STEM subject (Bundesministerium für Bildung, Wissenschaft und Forschung [BMBWF], 2022).

The IMST ("Innovations Make Schools Top") project (e,g., Krainer, 2021) supports these STEM middle schools in various focal points: networking, material development and evaluation. The material development supports the STEM middle schools in two ways. One aspect involves the examination and didactic preparation of current materials, intending to provide teachers with initial material for their lessons. Conversely, due to the lack of explicit teaching materials for the curriculum content so far, the second emphasis is placed on creating new teaching materials.

In existing materials, it is necessary to identify which characteristics should be addressed in STEM lessons and what role mathematics in particular can play. So far, characteristics of STEM (e.g., Thibaut et al., 2018) have been discussed independently of the role of mathematics (e.g., Just & Siller, 2022). The research objective that we are pursuing with this article is to assess and jointly discuss a selection of existing materials with regard to the characteristics of STEM lessons and the role of mathematics. From these results in joint consideration, endeavours for the creation of new materials will be derived.

THEORETICAL FRAMEWORK

Defining integrated STEM and frameworks

Moore et al. (2014) propose advocating for the integration of different STEM disciplines by focusing on complex application problems. Nevertheless, such a definition could unjustifiably exclude STEM education that is interdisciplinary but not strictly application-oriented (Kollosche & Schmölzer, in press).

Honey et al. (2014) present a definition of integrated STEM education which integrates application-oriented and non-application oriented approaches: "working in the context of complex phenomena or situations on tasks that require students to use knowledge and skills from multiple disciplines" (p. 52). In this definition, various forms of integration need consideration. Wang et al. (2011) distinguish between multidisciplinary and interdisciplinary integration: Multidisciplinarity arises from content and skills grounded in specific subjects, compelling students to establish connections between subjects encountered in different classes. Interdisciplinarity starts with a challenge that requires an understanding of both content and skills from multiple subjects. Additionally, Vasquez et al. (2013) introduce the transdisciplinary approach, wherein knowledge and skills from diverse disciplines are applied to address real-world issues.

In addition to efforts aimed at defining integrated STEM education, it is imperative to develop frameworks that effectively describe and evaluate STEM instruction. To advance the creation of valid assessments and protocols for researching integrated STEM teaching and learning, it becomes essential to articulate the characteristics of integrated STEM education in explicit detail.

Thibaut et al. (2018) condensed instructional practices through a systematic literature review and organized them within a theoretical framework. The theoretical framework identifies five fundamental principles associated with integrated STEM teaching. Table 1 provides an overview of the five principles, accompanied by a brief description of each one.

Inst. practices	**Description**
Integration of STEM content	"refers to the explicit assimilation of learning goals, content and practices from different STEM disciplines." (p. 8)
Problem-centred learning	"learning environments should involve students in authentic, open-ended, ill-structured, real-world problems to increase the meaningfulness of the content to be learned." (p. 8)
Inquiry-based learning	"learning environments that engage students in questioning, experiental learning and hands-on activities that allow them to discover new concepts and develop new understanding" (p. 8)
Design-based learning	"entails the use of open-ended, hands-on design challenges that provide students with the opportunity to not only learn about engineering design processes and engineering practices, but also deepen their understanding of disciplinary core ideas." (p. 8)
Cooperative learning	"students should get the opportunity to communicate and collaborate with each other to deepen their knowledge." (p. 8)

Table 1: Instructional practices in integrated STEM according to Thibaut et al. (2018)

The role of mathematics in integrated STEM

While the pinnacle of integration lies in combining all STEM disciplines equally to collaboratively work on real-world problems, such a comprehensive approach seems too demanding for most projects in STEM education. As a result, the specific contributions, potentials, limitations, and role of each subject within the framework of integrated STEM education can be explored. This analysis focuses particularly on the role of mathematics.

In their literature study, Just and Siller (2022) illustrate that mathematics is primarily used in many STEM activities as a calculation-oriented tool and for presenting results, leading to its perception only as a tool. However, fundamental mathematical concepts and methods often remain hidden, and the significance of mathematics for reflective decision-making processes is frequently overlooked. This is attributed, among other things, to the fact that science, technology, and engineering usually rely on mathematical models (Siller & Weigand, 2023). In integrated STEM lessons, mathematics could play a crucial role, serving as a platform where new mathematical concepts and methods are introduced or mathematical skills are expanded within an interdisciplinary context. These findings are taken as an opportunity to examine the existing role of mathematics.

The integrated STEM curriculum in Austria

The curriculum in Austria designates 11-15 extra hours for integrated STEM instruction across the four school years Each school has the autonomy to determine in which class and how many sessions of integrated STEM teaching should be conducted. The core subject-specific STEM principles primarily concentrate on an interconnected, interdisciplinary viewpoint. What holds paramount importance here is the understanding of broader contexts, encompassing the development of awareness regarding pivotal future issues. In Austria, the content is labelled as "interdisciplinary"; however, upon scrutinizing the curriculum, we can denote it as a transdisciplinary approach. Proficiencies from diverse disciplines should be utilized to address real-world challenges, such as environmental concerns, life cycles, or the realms of employment and digitization, as outlined in the curriculum (BMBWF, 2022). The emphasis on content and the reference to individual subjects are thus highly autonomous, left to the school or teacher.

METHODS

Existing materials

In 2022, the BMBWF commissioned IMST to review, further develop and create new teaching materials for the STEM subject for the STEM middle school pilot project as part of the work package for material development. At the beginning of the 2023/24 school year, material recommendations were made for the 2nd grade in the subject STEM, whereby existing materials were used, including a didactic commentary.

The materials are very heterogeneous in content as well as structure, ranging from material for one lesson to material for up to 29 lessons. The materials were chosen for this study for the pragmatic reason that this study, too, is part of the IMST project. However, they are also a good selection for this study, as they already underwent a quality check, where researchers and teachers from all STEM subjects discussed the suitability of the materials along pre-defined criteria such as age-adequacy, curricular relevance, teaching methods, interdisciplinarity, and inclusivity. These eleven material packages are the content of interest for the further analysis.

Analysis

Each of the eleven material packages was coded by two researchers, with the following elements annotated: number of lessons, instructional practices for STEM and the role of mathematics. The results of the coding procedure were summarised by consensus.

The instructional practices as well as the role of mathematics were categorized based on deductive category formation following Kuckartz (2018). We take the categories for the instructional practices from Thibaut's et al. (2018) framework. We classify the role of mathematics, following the categorization by Just and Siller (2022), into three categories: *new mathematical contents*, *mathematics used as an ancillary discipline*, or *no mathematical reference*. The category *nature of mathematics* is designated to STEM learning environments, where new mathematical concepts and methods are introduced, or mathematical skills are expanded within an interdisciplinary context. In contrast, the classification as an *ancillary discipline* is used if students apply pre-existing knowledge, utilizing mathematical terms and procedures as tools for problem-solving. STEM activities that lack mathematical references are categorized as to having *no mathematical reference*.

RESULTS

Table 2 displays the results from the instructional practices in integrated STEM analysed in the eleven materials. Each material is assigned a numerical study ID (in brackets), which is used for reference in the following text and tables.

ID	Material	Number of lessons	IC	PL	IL	DL	CL
[1]	Becoming Protectors of the Earth	16-29	x	x	x	x	x
[2]	The world's climate*	3-6	x	x	x		x
[3]	Paper airplane competition*	6	x	x	x	x	x
[4]	Plastic thought in circles*	6-24	x	x	x	x	x
[5]	ProtAct17	2-24	x	x	x		
[6]	Cleanly beaded*	4	x	x	x		x
[7]	Sound*	1-14	x	x	x		x
[8]	Material properties - A research trip*	9	x	x	x		x
[9]	Greenhouse effect in a drinking cup*	1-2	x	x	x		x
[10]	Our forests: Importance, threat, protection*	9-10	x	x			x
[11]	Virtual Lab	1-12	x	x	x		

Table 2: Instructional practices in integrated STEM materials. IC = Integration of STEM content; PL = Problem-centred learning; IL = Inquiry-based learning; DL = Design-based learning; CL = Cooperative learning; * Originally in German

Table 2 illustrates that existing teaching materials largely encompass the instructional practices of STEM education. It is noteworthy that only three out of the eleven materials exhibit elements of a design-based learning environment. The majority of the materials are in German (8 out of 11), with less emphasis on design-based approaches, and engineering is more in the background. This can be attributed to the fact that the English-language STEM explicitly includes “engineering”, a term that is not as explicitly expressed in the German MINT. Moreover, in the German-speaking school system, there is no emphasis on engineering in middle schools.

ID	New mathematical contents	Mathematics used as an ancillary discipline	No mathematical reference
[1]	x		
[2]		x	
[3]	x		
[4]			x
[5]		x	
[6]		x	
[7]		x	
[8]		x	
[9]		x	
[10]	x		
[11]		x	

Table 3: Role of mathematics in integrated STEM materials

The findings presented in Table 3 show that mathematical competencies are involved in ten out of the eleven reviewed STEM materials. Three of these materials emphasize mathematical concepts and methods within an integrated framework, considering both theoretical and applied aspects of mathematics. For instance, in material [3], methods for reliable data collection are discussed, new statistical measures for data analysis are introduced, and possibilities for presenting data are explored. Materials [1] and [10] comprise activities that focus on the mathematical modeling of real-life situations, the analysis of data, and the formulation of arguments based on them, leading to informed decision-making.

The seven STEM learning environments classified as *mathematics as an ancillary discipline* address real-world problems and projects, primarily examined from various scientific perspectives, with mathematics serving as a tool. Mathematical applications are limited to the collection of measurement data in experiments and the analysis and interpretation of diagrams, lacking an explicit connection to underlying mathematical concepts and methods. Upon closer examination of these materials, it becomes apparent that the role of mathematics can be strengthened through minor adjustments and mathematical excursus. Exemplary we have a closer look at material [2]. In this material, mathematics serves as an auxiliary discipline, requiring the interpretation of numbers from tables. The central theme revolves around weather and climate, providing an ideal opportunity to introduce negative integers. Focusing on the introduction of negative whole numbers aligns seamlessly with the mathematics curriculum for the second grade, creating a strong connection with traditional subjects. The content delves into the utilization of tables and climate diagrams, extending the understanding of negative numbers. This approach not only enriches mathematical concepts but also establishes a foundational understanding of negative integers within the context of weather patterns.

DISCUSSION

Table 2 and 3 highlight that teaching materials, within the framework of integrated STEM, encompass a variety of instructional practices aimed at promoting integrated STEM teaching. However, this analysis demonstrates that the current involvement of mathematics in STEM materials is often confined to an ancillary discipline.

The absence of specific learning goals of traditional school subjects and the broad scope of real-world applications give rise to the challenge that the STEM curriculum cannot establish clear connections between learning objectives and conditions in STEM and traditional subjects. If integrated STEM education is meant to improve competences in mathematics, then it should prioritize a heightened consideration and emphasis on the significance of mathematics. Integrated STEM education should intelligently build upon the content from traditional STEM subjects, where possible, intertwining with these subjects and only venturing into later content from traditional STEM subjects where necessary.

References

Bøe, M. V., Henriksen, E. K., Lyons, T., & Schreiner, C. (2011). Participation in science and technology: Young people's achievement-related choices in late-modern societies. *Studies in Science Education*, *47*(1), 37–72.

Bozkurt, A., Ucar, H., Durak, G., & Idin, S. (2019). The current state of the art in STEM research: A systematic review study. *Cypriot Journal of Educational Sciences*, *14*(3), 374–383.

Bundesministerium für Bildung, Wissenschaft und Forschung (BMBWF). (2022). *Lehrplan der Mathematik - Informatik - Naturwissenschaften -Technik Mittelschule (Kurzform MINT-MS) (im Schulversuch)* [STEM Curriculum].

Czerniak, C. M., Weber, W. B., Sandmann, A., & Ahern, J. (1999). A literature review of science and mathematics integration. *School Science and Mathematics*, *99*(8), 421–430.

Furner, J. M., & Kumar, D. D. (2007). The mathematics and science integration argument: A stand for teacher education. *Eurasia Journal of Mathematics, Science and Technology Education*, *3*(3), 185–189.

Honey, M., Pearson, G., & Schweingruber, H. (2014). *STEM integration in K-12 education*. National Academies Press.

Just, J., & Siller, H.-S. (2022). The role of mathematics in STEM secondary classrooms: A systematic literature review. *Education Sciences*, *12*(9), 629.

Kohen, Z., & Nitzan-Tamar, O. (2021). Excellence in mathematics in high school and the choice of STEM professions over significant periods of life. In M. Inprasitha, N. Changsri, & N. Boonsena (Eds.), *Pro. 44th Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 3, pp. 112–117). PME.

Kollosche, D., & Schmölzer, B. (in press) Grundlagen einer fächerübergreifenden MINT-Didaktik. *Didacticum*, *5*(2). https://didacticum.phst.at

Krainer, K. (2021). Implementation as interaction of research, practice, and policy. Considerations from the Austrian initiative IMST. *ZDM – The International Journal on Mathematics Education*, *53*(5), 1175–1187.

Kuckartz, U. (2014). *Qualitative text analysis: A guide to methods, practice & using software*. Sage.

Kurniati, E., Suwono, H., Ibrohim, I., Suryadi, A., & Saefi, M. (2022). International scientific collaboration and research topics on STEM education: A systematic review. *Eurasia Journal of Mathematics, Science and Technology Education*, *18*(4), 1–14.

Le Thi Thu, H., Tran, T., Trinh Thi Phuong, T., Le Thi Tuyet, T., Le Huy, H., & Vu Thi, T. (2021). Two decades of STEM education research in middle school: A bibliometrics analysis in Scopus database (2000–2020). *Education Sciences*, *11*(7), 353.

Moore, T. J., Stohlmann, M. S., Wang, H.-H., Tank, K. M., Glancy, A. W., & Roehrig, G. H. (2014). Implementation and integration of engineering in K-12 STEM education. In Ş. Purzer, J. Strobl, & M. E. Cardella (Eds.), *Engineering in Pre-College Settings* (pp. 35–60). Purdue University Press.

Nitzan-Tamar, O., & Kohen, Z. (2022). The effect of secondary mathematics on future choice in STEM professions. In C. Fernández, S. Llinares, Á. Gutiérrez Rodríguez, & N. Planas (Eds.), *Pro. of the 45th Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 3, pp. 243–250). PME.

OECD. (2018). *The future of education and skills: Education 2030*.

Reinhold, S., Holzberger, D., & Seidel, T. (2018). Encouraging a career in science: A research review of secondary schools' effects on students' STEM orientation. *Studies in Science Education*, *54*(1), 69–103.

Siller, H.-S., & Weigand, H.-G. (2023). Ohne Mathe geht es nicht. *Mathematik Lehren*, *2023*(237), 2–7.

Thibaut, L., Ceuppens, S., Loof, H. de, Meester, J. de, Goovaerts, L., Struyf, A., Boeve-de Pauw, J., Dehaene, W., Deprez, J., Cock, M. de, Hellinckx, L., Knipprath, H., Langie, G., Struyven, K., van de Velde, D., van Petegem, P., & Depaepe, F. (2018). Integrated STEM education: A systematic review of instructional practices in secondary education. *European Journal of STEM Education*, *3*(1), 02.

Vasquez, J. A., Comer, M., & Sneider, C. (2013). *STEM lesson essentials, grades 3-8: Integrating science, technology, engineering, and mathematics*. Heinemann.

Wang, H.-H., Moore, T. J., Roehrig, G. H., & Park, M. S. (2011). STEM integration: Teacher perceptions and practice. *Journal of Pre-College Engineering Education Research*, *1*(2), 1–13.

RELATIONALITY IN PRODUCTIVE STRUGGLE: A SOMALI ALGEBRA CONVERSATION

Susan Staats, Claire Halpert, Alyssa Kasahara, Emily Posson, and Fardus Ahmed

University of Minnesota, U.S.A.

This paper analyses relationality as a source of mathematical meaning during productive struggle in a multilingual, Somali and English algebra conversation. Relationality—meaningful interpretations based on interactions of multimodal dialogue, past language occurrences, mathematical writing, and learning environments—can take the form of conversational repetition. We show that the students' conversational repetition allowed them to express uncertainty in useful ways, exploring what it means to explain mathematically, and transforming the Somali meanings of words "add" and "write" in ways that enhanced their work towards algebraic generalization. Our analysis deepens the theoretical understanding of productive struggle when it involves uncertainty in explaining and sensemaking.

INTRODUCTION

Students engaged in productive struggle create mathematical meaning through interaction–always with their prior mathematical experiences; nearly always with institutionally-framed tasks; and usually through verbal interaction with other students and a teacher. Four student activities that may involve productive struggle are: 1) getting started; 2) carrying out a process; 3) uncertainty in explaining and sensemaking; 4) expressing misconceptions and errors (Warshauer, 2015, p. 385). Much productive struggle research is focused on teachers' strategies to sustain it (Warshauer, 2015). Relatively little research documents students' strategies for sustaining productive struggle when they are not immediately engaged with the teacher.

In this paper, we deepen the theoretical understanding of Warshauer's third dimension of productive struggle, uncertainty in explaining and sensemaking, through discourse analysis of two Somali-speaking undergraduates' algebraic problem-solving dialogue. Ambiguity or uncertainty is an understudied source of mathematical meaning (e.g., Foster, 2011). We examine the insight that "all mathematical meaning is relational" (Barwell, 2023, p. 537) with focus on relationality that arises through conversational repetition—revoicing and revising previous spoken comments or sentences in a written task (Staats, 2021). Our analysis of this multilingual conversation indicates that relationality through repetition calls forth uncertainty in explaining and sensemaking in two highly productive ways: 1) students' exploration of what it means to explain insights mathematically, and 2) their transformation of shared mathematical meanings.

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. 4, pp. 129–136). PME.

RELATIONALITY AND REPETITION IN MATHEMATICS DISCOURSE

This paper attends carefully to productive struggle that is noticeable in spoken, linguistic shifts as two students discuss an algebra task, but as Barwell points out, relationality is not carried exclusively by language: "meaning emerges from the interactions between students, the teacher and texts" (Barwell, 2023, p. 537). Forms of relationality that are important in this paper include: shifts in meaning compared to similar, recent comments from either a student partner or a researcher; spoken meanings that reference sentences or images in the task statement; and embodied relationality that occurs when students use collaborative writing, drawing or gesturing to create mathematical meaning.

Close analysis of speech, however, is the starting point in this paper to notice relationality as a source of mathematical meaning. We use the method proposed in Staats (2021) to identify grammatical repetitions, also known as conversational repetition or poetic structure. In this approach, a comment becomes a poetic echo of a previous comment if they share a repeated word or meaningful component of a word and some syntax. Meaning arises in the interplay of similarity and difference across a series of comments. These may be short, repeated comments or longer chunks of discourse, like a question-and-answer which is repeated several times. For example, two poetic sequences extracted from the selections in this paper include: *we found it…we wrote…we looked for it…we added…you are adding;* and later: *someone would be adding…someone would be adding...someone would be adding boxes…someone is adding on a row.* As one comment flows into the next, the similar words and syntax support a sense of relationality to past comments, but there are also changes in the phrasing, so that we end up in a different meaning-place from where we embarked.

These strings of relational adjustments can indicate important mathematical transitions. In pattern generalization tasks (Fig. 1), students' initial solutions are often numerical verification approaches that they sometimes extend through "empirical re-conceptualization" into a stronger deductive or structural generalization (Ellis, Lockwood, & Ozaltun-Celik, 2022). Conversational repetition can assist this process by allowing speakers to adjust their recent framing of a task in relation to newly-noticed mathematical features. We show that uncertainty—sometimes grammatically-encoded and sometimes emerging semantically across repetitions—is a linguistic resource for sensemaking and explaining during productive struggle.

METHOD

Two Somali-speaking undergraduates, who we call Leylo and Raana, worked for about 35 minutes in a paid, out-of-class, video and audio recorded setting, on the perimeter generalization task shown in Figure 1. Outside of this setting, the two are personal friends to each other. The full task (not shown here) was translated from English into Somali by other Somali-speaking undergraduates.

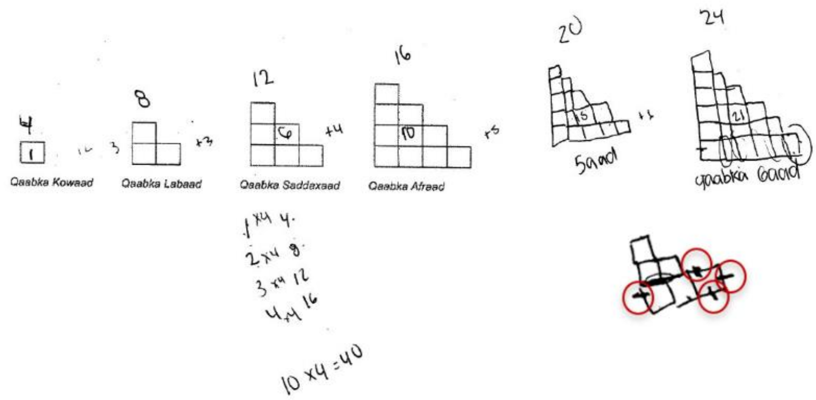


Figure 1: Perimeter task; two mathematical properties written lower on the page.

The recordings were translated through the linguistics-based interlinear morphemic glossing method described by Edmonds-Wathen (2019) that accounts for the semantic and grammatical contribution of all the morphemes, the meaningful components of an utterance, as in Leylo's comment:

Oh, row aa la darayaa!	*Verbatim transcription*
Oh row aa la dar-ay-aa	*Separation into morphemes*
Oh row LOC ISP add-PROG-PRES	*Abbreviated gloss / morphemic analysis*
Oh, someone is adding on a row!	*Free translation into English*

Morphemic glossing abbreviations accurately portray the agency and action within Leylo's comment, that an unspecified actor *la* (ISP, impersonal subject pronoun) is in the process of adding (PROG, progressive '-ing' sense) a row in a location (LOC, locative marker) in the present moment (PRES, present tense). Publication constraints do not allow presentation of the fully glossed dialogue, but our discussion of repetition and relationality is supported by this detailed analysis for the 35-minute conversation.

DIALOGUE AND DISCUSSION

After about 20 minutes, Leylo and Raana had noticed that the perimeter can be calculated using the calculation block shown in Fig. 1, and that the perimeter "goes up by four" as one moves from one case to the next. They shared these results in English with the first author, who confirmed them. The first author then noted that their method involved a "four, a multiplication, and a number that is changing," and suggested that they could further explain it in relation to the "shapes" of the cases. Before this conversation, Leylo and Raana felt certain of their method, but afterwards, they knew there was uncertainty associated with what a successful or complete explanation of their method might be. Subsequent work, including selections 1 and 2 below[1], resulted in the visual justification for "going up by four" shown in Fig. 1.

Productive uncertainty in explaining

In Selection 1, Leylo with support from Raana developed a question-and-answer pattern of talk that was repeated three times, each represented as a stanza (starting at

1.1.1, 1.2.1, and 1.3.1). Leylo's opening comment *Sidee ku helikartaa*, was a shortened form of a written task question that the two had answered in writing earlier, *Sidee ku helikartaa wareega qaabka X-add / How are you able to find the perimeter of the Xth case*? Leylo and Raana seem to endorse the task question as a way to develop a valid mathematical answer as a relational guide towards the act of doing mathematics. They answered the revoiced question at 1.1.1 with repetitions of *uu helnay / we found it* and *waa qornay / we wrote*, past tense verbs that convey a sense of factual reality, similar to their English translations. Most of the action in stanza 1 was portrayed as collaborative and completed work, with pronouns translating to *we*, but Leylo briefly used the impersonal pronoun *la* (later, in 1.3.1, *lo*) / *someone* or *one*, referencing an unspecified person who hadn't yet noticed any of the task properties.

Further, these repetitions of the verbs *hel- / found* flowing into *qor- / write* were instances of relationality, but they also referenced earlier, embodied relationality when the two spent minutes collaboratively writing sentences justifying their multiplication method, with one dictating while the other wrote, passing the paper back and forth, editing the other's sentences, with extreme rigor and deliberation. When Leylo and Raana said *waa qornay / we wrote*, the only thing that they had collaboratively written at that point were sentences explaining their method. It seems likely, then, that *qor- / write* referred to the shared actions of writing sentences, and that selection 1, stanza 1 expressed the sense that they were certain that their method was correct.

Selection 1, in three stanzas

[1.1.1] L: Ok, so number aa <u>la</u> bedel<u>ayaa</u>. <u>Sidee ku helikartaa</u>?

Ok, so someone is changing the number. How are you able to find (it)?

[1.1.2] **L: Because anaga saa uu <u>helnay</u>,** *(points at 1st case)*

Because thus we found it,

[1.1.3] **L: <u>waa qornay</u>.** *(sweeps across the calculations)*

we wrote (it).

[1.1.4] **R: <u>Waa qornay</u>.**

We wrote (it).

[1.2.1] L: So hadda ogin inaa <u>la</u> <u>qorikaro</u> sas *(sweeps calculations),* **<u>sidee ku</u> <u>helikartaa</u> qaabka** *(points at 5th, 6th cases, and below)* **ku so xigo?**

So if (one) didn't know that one is able to write like that, how are you able to find the next case?

[1.2.2] **L: So anaga waxaa <u>raadinay</u>, se like,** *(touching the 5th case)* **whenever you like**

So we looked for it, how like, whenever you like,

[1.3.1] L: Sidee lo raadiyo marka maxaa ku darnay?

How might someone look for it when we added on?

[1.3.2] L: **So halkan haddad sawirtid,** *(touches 1st case, the 4 above, and then the 1st case again)*

So here, if you would draw (it),

[1.3.3] L: **Halkaan aad ku daraysaa,**

Here, you are adding on**,**

[1.3.4] L: **Halkaan aad ku daraysaa**

Here, you are adding on**,**

[1.3.5] L: **Lakin wuxuu kacay sideed** *(touches the 8 above the 2nd case)*

But it rose to eight.

In stanzas 2 and 3, the *sidee / how* question-and-answer sequence was repeated with verbs that express both a richer sense of uncertainty compared to stanza 1, and productive struggle to articulate a new mathematical justification. In contrast to stanza 1—the factual voicing of *bedelayaa / is changing* (present tense) and of *we found it…we wrote*—in stanza 2, Leylo used the final sound *-o* in *qorikaro / able to write* which expresses an indefiniteness that is grammatically required by the negative state of someone not knowing in *ogin* (1.2.1). Both final *-o* and final *-tid* express this "irrealis" or uncertain state of a verb, and often take a translation such as *could*, *would* or *might* (1.2.1, 1.3.1, 1.3.2, 2.1.1, 2.1.3, 2.2.1).

Arguably, working towards a deductive generalization requires an epistemological stance of uncertainty or delayed belief towards what might be true mathematically. Leylo and Raana's three question/answer repetitions expressed this new uncertainty through irrealis verb markers and an impersonal, not-yet-knowing actor *la*. Through repetition, they associated this indefinite mathematical person with specific mathematical investigations that have potential to inspire explanation: *looking*, *drawing*, and *adding on* (1.2.2, 1.3.2, 1.3.3, 1.3.4). Selection 1 demonstrates repetition as a way to establish a stance of productive uncertainty towards the act of explaining.

Productive uncertainty in sensemaking

Importantly, the selection 1 comment regarding *dar- / add*, *How might someone look for it when we were adding it on?* (1.3.1) was the conversation's first mention of addition. Similar to English, the Somali verb *dar-* can mean either numerical *adding* or physical *appending*, *attaching*. At 1.3.1, we can't tell which meaning was in play, but in the passages after selection 1 and before selection 2, *dar-* emphasized the sense of adding as a numerical operation, with comments like "*this one is adding three,"* while writing + *3* on the task images. The Somali verb *qor- / write* has slightly different boundaries than English *write*, because *qor-* can mean either *writing* sentences or *drawing* boxes, but in English, one would rarely say that we *write boxes*.

Previously, Leylo had drawn the 5th and 6th cases by working with columns rather than rows, attaching a left-side column to the 4th and 5th cases (mentioned in 2.1.2). But at 2.2.2, the pair realized that they could also build the next case by attaching a row on the bottom of the previous case. This breakthrough of drawing rows rather than columns seemed to offer an easier way to generalize, that a new row always yields four sides in the perimeter that were not previously counted (Fig. 1). The "reversal" repetition (Staats, 2021) from 2.1.3 to 2.2.2 brings "rows" in front of *daraya* instead of "boxes" after *daraayo*, in Somali grammar, marking "rows" as important, new information, and removing the uncertainty of the irrealis marker *-o*. Conversational repetition allowed Leylo and Raana to re-specify the meaning potentials of *dar-* towards *appending* rather than *adding*, and *qor-* towards *drawing* rather than *writing*.

Selection 2, in two stanzas

[2.1.1] L: But I think this is easier, mid <u>la</u> explaingaraynkar<u>o</u>,

But I think this is easier, which someone would be able to explain,

[2.1.2] L: because <u>waa qoray</u> boxes,

because (I) wrote boxes

[2.1.3] L: lakin I think <u>marka lagu</u> isku <u>darayo</u> the boxes,

but I think when someone would be adding the boxes onto itself,

[2.1.4] L: Why, I don't know, which is what I am trying to figure out

[2.2.1] L: <u>Marka habad lagu darayo</u>,

When someone would be adding one on

[2.2.2] L: <u>Oh, row aa la daraya</u>! *(tracing bottom of 3rd and 4th case)*

Oh, someone is adding on a row!

[2.2.3] R: <u>Oh, the rows</u>. Haa!

Oh, the rows. Yes!

[2.2.4] R: Haa! Hadda firi -- I forgot what I was saying. *(holding paper near 1st case)*

Yes! Now look – I forgot what I was saying.

[2.2.5] L: Haa, <u>rows aa lagu daraya</u>, *(points to 1st case)* **<u>so every one</u>, <u>habadi</u>,**

Yes, someone is adding on the rows, so every one, this one here.

[2.2.6] **R: <u>habadi, habadi uu sii raca</u> –** *(R touches near the 1st case, then slightly to the right.*

this one here, this one here, must give it, continuing

[2.2.7] **L: <u>habadi row uu helayaa</u>** *(L touches the 3rd shape)*

This one here, the row is receiving it

[2.2.8] **R: <u>haa, the bottom, and then it goes like that</u> …**

(R continues L's sentence, while tracing across the all the cases' bottom rows)

yes, the bottom, and then it goes like that…

As the relational meaning framework suggests, words have a range of meanings that are always potentially available (Barwell, 2023). We can follow the sequential shift in meanings by reading an extracted, sequential poetic string of comments across selections 1 and 2:

waa qornay / we wrote (sentences)	1.1.3 and 1.1.4
ku darnay / we added on (numerical or figural is undeterminable)	1.3.1
waa qoray / I wrote boxes	2.1.2
daraayo the boxes / someone would be adding on the boxes	2.1.3
daraayo / someone would be adding	2.2.1
row aa la daraya! / someone is adding a row!	2.2.2

Each comment is discernibly a repetition of a previous comment through shared words or morphemes and syntax. This poetic string is a source of relational mathematical meaning, because each comment takes meaning from previous ones, sometimes stabilizing or continuing meanings and sometimes shifting or extending them. Across selections 1 and 2, these comments contributed, first, to a new stance on uncertainty in posing mathematical questions and seeking explanations, and then, in making ambiguous semantic meanings more specific at a moment of mathematical insight.

The last moments of selection 2 enact relationality through conversational repetition in creative and embodied ways. Raana and Leylo's spoken embodiment resulted in a sentence of generalized explanation from 2.2.5 to 2.2.8. Together they found a novel kind of poetic imagery to express the generalization, that one case *gives* a row and the next case *receives* it, both words expressing movement, transformation and connection but from opposing standpoints. And finally, their shared finger movements indicated generalization: *and then it goes like that*, with Raana's finger tracing towards infinity.

CONCLUSION

Mathematical productive struggle implies change across a series of events within a state of uncertainty or ambiguity. Our analysis shows students' productive struggle dialogue can usefully harness uncertainty or ambiguity through shifting grammatical and semantic interpretations of the mathematical task. The concept of "shifting" is important, that productive struggle emerges through the relationships across comments and actions, not within a logical parsing of a single sentence or action. Through conversational repetition, students can shift from certainty to uncertainty in a question-and-answer sequence, inviting new mathematical investigations. Repetition also allows speakers to sculpt ambiguous semantic meanings into foundations for new mathematical insights. Indeed, productive struggle research that attends closely to students' words and actions could usefully distinguish uncertainty (like irrealis verb forms) from ambiguity (like multivocalic words) (c.f. Foster, 2011). Through the relationality that is inherent in dialogue, even when a teacher is not present, students can sometimes support their own productive struggle, taking complex epistemological stances on what it means to know and to believe, mathematically.

NOTES

1. The numbering system is [selection.stanza.line], that (1.3.2) references selection 1, stanza 3, line 2. Underlining indicates comments of interest discussed in the paper. Indentation draws attention to stanza structure and repeated phrases, based on morphemic glossing. Words in parentheses were not spoken directly in Somali but create a smoother English translation.
2. This study was funded by a Grant-In-Aid of Research from the University of Minnesota Office of the Vice President for Research.

References

Barwell, R. (2023). Sourcing mathematical meaning as a dialogic process: Meaning-focused and language-focused repairs. *ZDM–Mathematics Education*, 55, 535-550.

Edmonds-Wathen, C. (2019). Linguistic methodologies for investigating and representing multiple languages in mathematics education research. *Research in Mathematics Education*, *21*(2), 119-134.

Ellis, A., Lockwood, E., & Ozaltun-Celik, A. (2022). Empirical re-conceptualization: From empirical generalizations to insight and understanding. *The Journal of Mathematical Behavior*, *65*, 100928.

Foster, C. (2011). Productive ambiguity in the learning of mathematics. *For the Learning of Mathematics*, *31*(2), 3-7.

Staats, S. (2021). Mathematical poetic structures: The sound shape of collaboration. *The Journal of Mathematical Behavior, 62*, 100846.

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

ADDITION AND SUBTRACTION PROFICIENCY INVOLVING NEGATIVE INTEGERS IN ZAMBIA

Shun Sudo[1], Koji Watanabe[2], and George Chileya[3]

[1]National Institute of Technology Hakodate College, Japan
[2]Miyazaki International University, Japan
[3]Ministry of Education Directorate of National Science Centre, Zambia

Zambia is recognised for its low academic proficiency in south-eastern Africa. Urgent attention is needed to develop basic arithmetic skills, as seen in children, resorting to methods such as drawing sticks to count for calculations such as 7 + 9, and only one in three correctly computing –4 – 2. This study reveals the calculation algorithms used by Zambian children in addition and subtraction, including operations with negative integers. When providing incorrect answers, children associate calculations with those they can already perform correctly. This insight highlights the importance of developing instructional strategies that build upon existing abilities to address the pressing need for enhancing basic arithmetic proficiency in Zambia.

INTRODUCTION

The Sustainable Development Goal 4 (SDG 4) focuses on education, encompassing ten targets, one of which emphasises the urgent need to ensure that “all young people and a substantial proportion of adults, both men and women, achieve literacy and numeracy”. This highlights the importance of fostering basic numeracy skills to achieve high-quality education, particularly in developing countries.

In Southeast Africa, a specific educational assessment known as the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ) reveals that Zambia ranks lowest in academic proficiency among the countries (Musonda & Kaba, 2011). Regarding the reality in Zambia, for instance, in a calculation such as 7 + 9, many primary school students tended to draw seven lines, subsequently add nine more lines, and finally count the total number of lines drawn. Reports indicate that methods involving “drawing lines” and “counting using fingers” dominate the approach to calculations (JICA, 2021). Furthermore, when it comes to addition and subtraction involving negative integers, such as in the case of – 4 – 2, reports indicate that 8th and 9th-grade students in Zambia use a number line to calculate; however, the accuracy rate for –4 – 2 is approximately 33%. This means that only one out of three students could correctly perform this calculation (Sudo et al., 2019). These findings illustrate the lack of basic computational skills, which poses a significant challenge to achieving SDG 4.

Several patterns of addition and subtraction involve negative numbers. When both variables a and b are positive integers, their sum, denoted by a + b, is always a positive integer, regardless of their relative sizes. However, regarding variables a and b, 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. 4, pp. 137–144). PME.

result can be positive or negative depending on the relative sizes of a and b. Similarly, concerning negative numbers, –a + b (where both a and b are positive integers), the sum can be positive or negative based on the relationship between a and b. For –a – b, the difference is always a negative integer, regardless of the sizes of a and b. In summary, there were four scenarios: a + b, a – b, a + b, and –a – b. For a – b and –a + b, the outcome depends on whether variable a is greater than, equal to, or less than variable b. Additionally, viewing –a – b as –a + (–b) allows us to consider the subtraction involving negative numbers as the addition of negative integers. Considering the relative sizes of variables a and b from this perspective further increases the complexity of the calculation patterns. Thus, calculations involving negative integers are more intricate than those in which the sum and difference remain positive.

There is a substantial body of research on calculations involving negative integers (Sahat et al., 2018; Salsabila et al., 2022). Many of these studies focused on teaching addition and subtraction using number lines (Fuadiah et al., 2019). Although Sudo et al. (2019) reported that children in Zambia also used number lines for calculations, their mastery of this method was insufficient. Sudo et al. (2022) conducted a similar investigation in Uganda and reported a correct answer rate (CAR) of approximately 21% for –4 – 2 calculations. They highlighted that similar to Zambia, the mastery of calculations involving negative integers is insufficient in Uganda.

Sudo et al. (2019) and Sudo et al. (2022) conducted a calculation test comprising 24 questions that differentiated the sizes of variables a and b in the expressions a + b and a + (–b). In this study, we aimed to investigate the actual computational abilities of children. The results revealed low levels of computational proficiency. However, despite not arriving at correct answers, the specific calculation processes employed by children have not been explicitly elucidated. For instance, in Zambia, where only approximately 33% of the children provided the correct answer for –4 – 2, it is possible that the remaining 67% of the children employed some form of calculation method, leading to incorrect answers. This implies that even in cases where the correct answer is not achieved, children may utilise a computational algorithm. It is essential not only to address instances of incorrect calculations but also to uncover the specific processes that children employ, even when arriving at an incorrect answer. This point of view has not yet been examined in earlier research.

In this study, we aimed to elucidate the computational algorithms employed by Zambian children in their incorrect answers. Between 2017 and 2019, a survey was carried out in eight secondary schools in Zambia. The survey utilized a calculation test comprised of 24 questions, which was conducted by Sudo et al. (2019). It is important to note that Sudo et al. (2019) utilised data specifically from the 2017 survey. In our analysis, we first determined the difficulty levels of the 24 test questions using both classical test theory and item response theory. Subsequently, we aimed to uncover the typical incorrect answers of the children and reveal the unique computational algorithms employed by these children that lead to incorrect answers. This comprehensive approach included data from the entire period (2017–2019) covered by the original study, providing a more extensive perspective on the children's computational strategies.

METHODS

This analysis focuses on data collected from a survey of 8th and 9th-grade students who completed learning addition and subtraction involving negative integers from 2017 to 2019 in eight schools. The total sample size was 971 children (396 in 2017, 204 in 2018, and 371 in 2019). The test consisted of 24 questions related to addition and subtraction involving negative integers, as presented in columns A and B of Table 1. For this test, positive integers a and b were used, covering 12 types ranging from a + b to –a – (–b). These types were created by setting the addends or subtrahends as positive or negative integers and forming combinations with positive and negative integers as augments or minimums. Each type was further divided into Types A and B, distinguishing between a > b and a < b. For example, 4 + 2 and 2 + 6 were considered the same type (Type (1)), and both Types A and B were included. In the six types (Types (2), (3), (6), (7), (9), and (12)), the signs of the sum or difference differed between Types A and B. This characteristic adds complexity to the test, which is one of the reasons why addition and subtraction involving negative integers are considered challenging. Additionally, to simplify the calculations, the values of a and b were even, ensuring that the sum or difference in each question was a single digit.

RESULTS

The analysis of the difficulty of the 24 questions

The difficulty of the 24 questions was estimated using both classical test theory and item response theory; the results are presented in Table 1. In classical test theory, Cronbach's alpha reliability coefficient was calculated and yielded a value of 0.854, indicating sufficient reliability. The difficulty level of each item was assessed based on the CAR. To estimate difficulty (Dffclt) using item response theory, the eigenvalues of the tetrachoric correlation matrix for all 24 questions were calculated and their respective contribution rates were examined. The contribution rates for the first three eigenvalues were 36%, 10%, and 9%, respectively, with the first eigenvalue being noticeably larger. This indicated that the test satisfied the assumption of unidimensionality. The Rasch model, which considers only item difficulty as a parameter, was employed to estimate item difficulty using item response theory. The histogram of test scores estimated using the Rasch model is illustrated in Figure 1, with mean and standard deviation values of –0.003 and 1.024, respectively.

Table 1: Test questions and their difficulty

Type	Type A	CAR	Dffclt	Type B	CAR	Dffclt
(1) a + b	4 + 2	97%	–3.897	2 + 6	97%	–4.039

(2) a – b	6 – 4	93%	–3.043	4 – 8	48%	0.092
(3) –a + b	–8 + 4	67%	–0.934	–6 + 8	51%	–0.101
(4) –a – b	–4 – 2	35%	0.755	–2 – 6	33%	0.855
(5) a + (+b)	4 + (+2)	82%	–1.846	2 + (+6)	83%	–1.936
(6) a + (–b)	8 + (–2)	54%	–0.246	4 + (–8)	60%	–0.550
(7) –a + (+b)	–4 + (+2)	58%	–0.435	–4 + (+8)	52%	–0.147
(8) –a + (–b)	–6 + (–2)	52%	–0.142	–2 + (–6)	54%	–0.251
(9) a – (+b)	8 – (+2)	58%	–0.429	4 – (+6)	42%	0.394
(10) a – (–b)	6 – (–2)	37%	0.619	2 – (–4)	31%	0.969
(11) –a – (+b)	–4 – (+2)	30%	1.043	–2 – (+6)	27%	1.215
(12) –a – (–b)	–6 – (–2)	47%	0.129	–6 – (–8)	39%	0.514

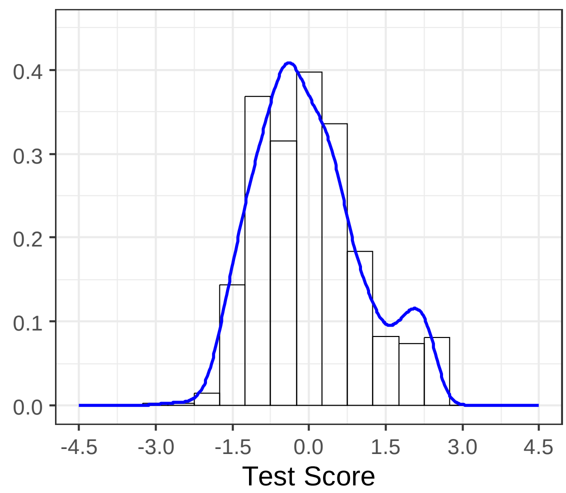


Figure 1: Histogram of the test score

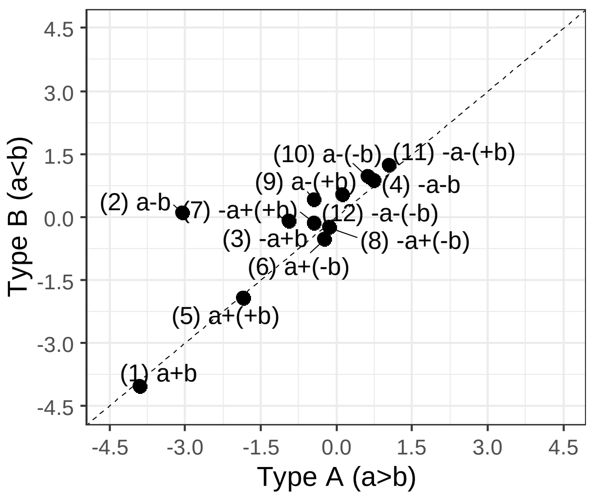


Figure 2: Scatter plot of the difficulty by type

The most challenging question, –2 – (+6), has a difficulty level of 1.215. This means that a child with a test score of 1.215 had a 50% probability of answering the question correctly. Considering the standard deviation of 1.024 for the test scores, this difficulty level is considered high. The CAR appeared to be challenging, with only one in three children answering correctly. Conversely, the question with the lowest difficulty level, 2 + 6, had a difficulty rating of –4.039. This indicates the difficulty level at which almost all the examinees are likely to answer correctly. The correct response rate was 97%, indicating that nearly all participants responded to this question correctly.

Figure 2 illustrates a scatter plot of the difficulty levels for Types A and B across the 12 types. In Type (2), it is evident that the difficulty level for Type B is noticeably higher than that for Type A. Conversely, for the remaining types, although the difficulty of Type B is slightly higher, there is no prominent difference, similar to that observed in Type (2). Among the other three types that do not include parentheses, Type (1) involves addition without negative integers and has the lowest difficulty level. For Type (3), the difficulty levels for Types A and B were –0.934 and –0.101, respectively. In Type A, one in three participants made a mistake, whereas in Type B, approximately half of the participants provided incorrect answers. In Type (4), both Types A and B had high difficulty levels of 0.755 and 0.855, respectively, making it a challenging question, with only one in three participants answering correctly.

The provided set of 24 questions represents fundamental calculations. Specifically, among the questions that did not include parentheses, despite covering the basics of addition and subtraction involving negative integers, there were instances in which only one in three children answered correctly. This suggests a situation where mastery of addition and subtraction involving negative integers cannot be deemed sufficient.

Analysing the wrong answers

For example, in Type (4), where only one in three children appears to calculate correctly, more participants may provide incorrect answers. This suggests that specific calculation algorithms may lead to incorrect responses. In this context, we examine the nature of these errors. It is worth noting that the analysis focuses on Types (1) through (4) without parentheses because including calculations with parentheses from Types (5) to (12) complicates the analysis.

For Types (1)–(4), the top three answers with the highest frequency for both Types A and B among the eight questions were examined. Across all eight questions, the correct answer (CA) was identified among the top three. The answer with the highest percentage, considered the most frequent wrong answer, was labelled WA1 (wrong answer), followed by WA2 with the second highest percentage, and the remaining answers were categorised as WA3. The percentages were calculated and the results are presented in Table 2.

For Type (1), specifically question (1A) 4 + 2, CA was 6, whereas WA1 and WA2 were 8 and –6, respectively. The correct response rates for (1A), (1B), and (2A) were all above 90%, indicating that the majority of participants could accurately answer these questions. Conversely, examining question (4A) –(–4) –2, the percentage of WA1 (–2) was 42% higher than the correct answer of –6, highlighting a notable misconception among participants. Similarly, (2B) and (4B) exhibit high percentages of WA1 at 37% and 28%, respectively.

Table 2: Classification and proportion of answers of Types (1) to (4)

Type	Type A/B	Dffclt	CA	WA1	WA2	WA3
(1) a + b	(1A) 4 + 2	–3.897	6 (97%)	8 (1%)	–6 (1%)	(1%)
	(1B) 2 + 6	–4.039	8 (97%)	4 (1%)	–8 (1%)	(1%)
(2) a – b	(2A) 6 – 4	–3.043	2 (93%)	–2 (3%)	10 (2%)	(3%)

	(2B) 4 – 8	0.092	–4 (47%)	4 (37%)	12 (7%)	(8%)
(3) –a + b	(3A) –8 + 4	–0.934	–4 (67%)	–12 (22%)	12 (4%)	(7%)
	(3B) –6 + 8	–0.101	2 (51%)	–14 (19%)	–2 (14%)	(15%)
(4) –a – b	(4A) –4 – 2	0.755	–6 (35%)	–2 (42%)	2 (12%)	(11%)
	(4B) –2 – 6	0.855	–8 (33%)	4 (28%)	–4 (20%)	(18%)

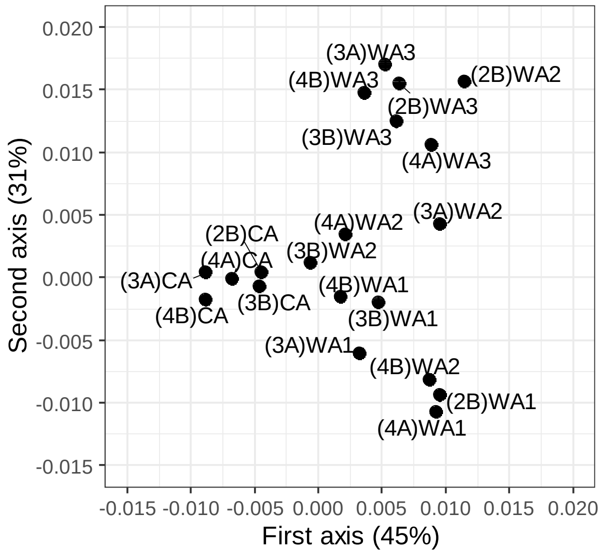


Figure 3: Result of correspondence analysis

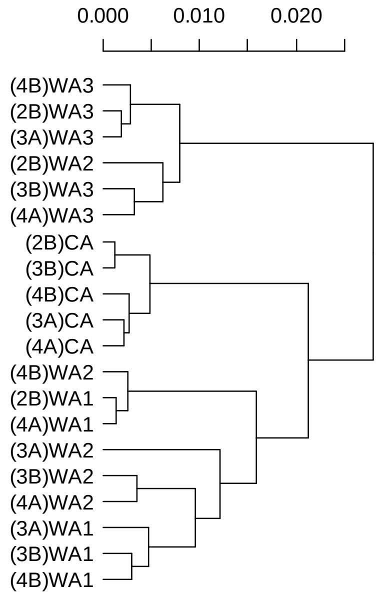


Figure 4: Result of cluster analysis

Therefore, for the five questions (2B), (3A), (3B), (4A), and (4B), correspondence analysis was conducted to reduce the information to a two-dimensional plane for the CA, WA1, WA2, and WA3 responses. As illustrated in Figure 3, the contribution rates for the first and second axes were 45% and 31%, respectively. Subsequently, cluster analysis was performed on this coordinate plane, as depicted in Figure 3, and the responses were grouped into clusters at a threshold of 0.018, as depicted in Figure 4. The upper cluster represents a collection of WA3 responses other than those of WA1 and WA2. The middle cluster consisted of correct answers (CA). The lower cluster comprises a collection of frequently observed incorrect answers, specifically for WA1 and WA2. Thus, the distinctive characteristics of each cluster suggest the presence of unique calculation algorithms that led to incorrect answers, particularly in the lower cluster.

First, (1A), (1B), and (2A) exhibited high correct response rates and low difficulty, indicating that the participants could answer these questions accurately. Assuming this premise, for the other five questions, WA1 and WA2 included numbers obtained through the addition or subtraction of certain numbers within the given calculations. For example, for WA1 in (2B), 4 can be obtained from 8 to 4. Similarly, in (3B), WA1, where –14 is present, can be obtained through 6 + 8, and WA2, where –2 is present, can be obtained through 8 – 6. This suggests an algorithm in which, even if the correct answer is not achieved, participants reinterpret the given mathematical expressions into a form that can already be calculated, similar to (1A), (1B), and (2A). Additionally, (3A), (3B), (4A), and (4B) all involve negative integers as either addends or minute ends. For instance, in (3A), WA1, which is –12, can be obtained through –(8 + 4), (3B) WA1 can be obtained through –(6 + 8), (4A) WA1 can be obtained through –(4 – 2), and (4B) WA2 can be obtained through –(6 – 4). This reveals an algorithm where participants write the "–" symbol without considering its meaning as indicating negative integers. They then calculate based on subsequent operations if they are already familiar with the calculation.

DISCUSSION

Using the 24-question test, we determined the difficulty of each question and revealed the unique calculation algorithms that led to incorrect responses among the children. The results indicate that apart from the three addition and subtraction questions, Type (1) Type A, B, and Type (2) Type A, children have not sufficiently mastered addition and subtraction, which are elementary school-level topics. Particularly noteworthy is the insufficient proficiency in basic calculations, even for questions that do not involve parentheses.

Therefore, we analysed incorrect answers focusing on the four types (Types (1), (2), (3), and (4)) of the eight questions, in addition to subtraction involving negative integers, excluding parentheses. Two distinct calculation algorithms emerged among the children, leading to incorrect answers. The first algorithm involves reinterpreting the calculations into a form that can already be solved accurately, similar to Types (1) a + b (a > b and a < b) or (2) a – b (a > b). The second algorithm pertains to calculations in which negative integers are either addends or minuends, such as Type (3) –a + b and Type (4) –a – b. In these cases, the participants wrote the "–" symbol as is and proceeded to calculate it based on a method with which they were already familiar. For example, –a + b is calculated as –(a + b) and –a – b as –(a – b). Thus, even in the case of incorrect answers, it was evident that the Zambian children associated their calculations with those they could already perform.

If, hypothetically, children who employ their unique calculation algorithms based on calculations they already know arrive at correct answers, then the accuracy rates for all eight questions in Table 2 would exceed 80%. Even in cases of incorrect answers, given that children utilise some form of a calculation algorithm, it can be reasonably expected that a thorough understanding of the calculation algorithm leading to correct answers would significantly enhance the likelihood of arriving at correct solutions.

Based on the unique calculation algorithms identified for children, where –a + b is computed as –(a + b) and –a – b is computed as –(a – b), it appears that there is a lack of understanding of the symbol "–" representing negative numbers. In other words, there is insufficient understanding of the existence and representation of negative numbers before engaging in addition and subtraction involving negative numbers. Therefore, future educational efforts should focus on fostering a comprehensive understanding of negative numbers. While this study specifically targeted children's calculations involving negative integers, future initiatives should aim to demonstrate children's actual understanding of negative numbers. Developing specific teaching methods based on the results of this study and a fundamental understanding of negative numbers is crucial. By conducting such research and developing effective teaching strategies, we can contribute to addressing the global challenges of fostering fundamental arithmetic skills.

Acknowledgement

This study was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI, grant numbers 18K13186 and 20H01689.

References

Fuadiah, N. F., Suryadi, D., & Turmudi. (2019). Teaching and Learning Activities in Classroom and Their Impact on Student Misunderstanding: A Case Study on Negative Integers. *International Journal of Instruction*, 12(1), pp. 407-424. https://doi.org/10.29333/iji.2019.12127a

JICA (Japan International Cooperation Agency). (2021). *Project Research for Analysis of Challenge and Solutions for Improving Numeracy Competence in Zambia Project Completion Report*. JICA. https://openjicareport.jica.go.jp/pdf/12362984.pdf

Musonda, B., & Kaba, A. (2011). *The SACMEQ III Project in Zambia: A Study of the Conditions of Schooling and the Quality of Education*. SACMEQ (Southern and Eastern Africa Consortium for Monitoring Educational Quality). https://www.sacmeq.org/sites/default/files/sacmeq/reports/sacmeq-iii/national-reports/s3_zambia_final.pdf

Sahat, N., Tengah, K. A., & Prahmana, R. C. I. (2018). The Teaching and Learning of Addition and Subtraction of Integers Through Manipulative in Brunei Darussalam. *Journal of Physics: Conference Series*, 1088, pp. 012024. https://doi.org/:10.1088/1742-6596/1088/1/012024

Salsabila, I., Amir, M. F., & Wardana, M. D. K. (2022). A Learning Trajectory of Integer Addition and Subtraction using the Kempreng Game Context. *Journal of Elementology*, 8(2), pp. 556-571. https:/doi.org/10.29408/jel.v8i2.5541

Sudo, S., Chileya, G., & Moonga, A. (2019). Metaphors and Errors in Calculation Skills of Integers in Zambian Learners. *Zambia Journal of Teacher Professional Growth (ZJTPG)*, 5(2), pp. 1-15.

Sudo, S., Kume, A., Bagenda, D., & Fujino, Y. (2022). Causes of Low Mathematics Achievement in Uganda and Exploiting ICT-based Tools in Response. *2022 International STEM Education Conference (iSTEM-Ed 2022)*. https://org/:10.1109/iSTEM-Ed55321.2022.9920900

CREATING A SENSE OF BELONGING IN THE ELEMENTARY MATHEMATICS CLASSROOM: RESPONDING TO (SOME OF) PAOLA VALERO'S 2023 PME PLENARY

Eva Thanheiser, Molly Robinson, Amanda Sugimoto, Simon Han,
Portland State University

Courtney Koestler, Mathew Felton-Koestler
Ohio University

Mathematics classrooms need to be spaces where each student experiences a sense of belonging, but what does this look like in an elementary mathematics classroom? To examine this issue, we designed lessons that allowed all students to see themselves and their classmates in the data they examined and thus learn mathematics while learning about themselves. We videotaped these lessons and analysed them using the construct of Belonging. We found that allowing students to explore ideas about themselves both allowed them to experience a sense of belonging as well as engage in the mathematics and contextual goals of the activity. We identified target teaching and learning practices to achieve these goals.

RATIONALE

In last year's PME plenary Paola Valero shared that "mathematics education does something with/in "learners", that goes beyond transformations in their cognition and thinking around mathematical notions and procedures," (Valero, 2023, p. 56). More specifically that "mathematics education leaves marks in children and learners" (Valero, 2023, p. 57) which can sometimes leave children alienated from mathematics and from the classroom community. For example, Janine, a special education teacher in the United States who identifies as a Black woman, shared her experience as a mathematics learner during her childhood.

> [In elementary school I am] starting off as a young child thinking I can do anything. Math was not a problem. … [Then] I moved from a school that was predominantly BIPOC, black, inner city … to a private school, Catholic school, where I was, like, probably a few of the black students. … So here I am confused, I am ostracized, I am not called on. I am… the expectations for me... [choked up] Sorry. Obviously, that's a painful part, but that is where the expectations, even though I was in a private school, were very limited for me and I wasn't called upon and sat in the back of the room. This is in the late '80s/'90s.

Janine carries marks from her mathematics experience that go beyond mathematics learning. She did not feel a sense of belonging, and these experiences clearly excluded her from the community in the mathematics classroom. In this paper we examine how focusing on a sense of belonging in the mathematics classroom might have allowed her and more children to feel part of mathematics rather than being harmed by it. Our research questions are: How can we create a sense of belonging for children in the elementary mathematics classroom? More specifically, can using data about

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. 4, pp. 145–152). PME.

themselves create a sense of belonging for the children in an elementary mathematics classroom?

FRAMING AND LITERATURE REVIEW: BELONGING IN THE MATHEMATICS CLASSROOM

Belonging has been identified as a main indicator of success, persistence, and well-being (Cwik & Singh, 2022). Belonging increases engagement and self-efficacy, which leads to success (Bandura, 1977, Trujillo & Tanner, 2014). Students who are confident they belong and feel valued by their instructors and peers can engage more fully in learning (Freire, 2018).

In this paper we define belonging as “a basic human need, a fundamental motivation, sufficient to drive behaviors and perceptions. Its satisfaction leads to positive gains such as happiness, elation, well-being, achievement, and optimal functioning” (Strayhorn, 2019, p. 9). More specifically we follow Watson’s (2024) **five factors of belonging** in mathematics which are: “You see me, I see myself, Others see me, I am valued, and I can grow” (retrieved January 2024, Crystal M. Watson). Watson’s five factors align with Strayhorn’s sense of belonging as a perceived sense of social support and feelings of connectedness. Minoritized students (like Janine above) have been shown to experience less of a sense of belonging (Gopalan & Brady, 2019) in general and specifically in the mathematics classroom.

To increase a sense of belonging for all students we build on Teaching Mathematics About/With/For Social Justice, extending from the work of the authors from the Middle School Mathematics Lessons to Explore, Understand , and Respond to Social Injustice (Conway et al, 2023) and from The Benjamin Banneker Association’s statement on Teaching Math for Social Justice. We focus on teaching mathematics WITH social justice (TMWSJ), which includes designing interactions within and beyond the classroom that attend to participation, status, and positioning. We examine TMWSJ as implementing **teaching practices** that allow students to engage in mathematics learning while learning about themselves and the world addressing participation, status, and positioning (developed based on Bartell et al., 2017). These practices include:

1. Maintain a focus on real-world reasoning.
2. Maintain a focus on mathematical reasoning.
3. Facilitating socio-political discourse.
4. Facilitating mathematical discourse.
5. Eliciting and using evidence of student thinking.

In addition we draw on the Social Justice Standards (Learning for Justice, 2016) which focus on Identity (Relate how identity has many characteristics and affects relationships within and beyond the classroom), Diversity (Develop respectful ways to discuss similarities and differences with others and begin to think about how diversity affects relationships within the classroom and beyond), Justice (Understand the difference between personal stereotypes and systemic discrimination, and explore how privilege impacts discrimination and justice), and Action (Move students from prejudice reduction to collective action). We focus on **Diversity** in this paper.

Finally, to create belonging in the mathematics classroom the definition of mathematics must be broadened from a narrow view of mathematics as an abstract body of knowledge/ideas, the organization of that into systems and structures, and a set of methods for reaching conclusions to, instead, mathematics as a verb (not a noun), a human activity, part of one's identity (Thanheiser, 2023). Thus, students must be allowed to share their ideas and follow through with those and see themselves and others in the mathematics. We focus on two learning practices: Students seeing/learning about themselves and others in the class, and Students sharing ideas.

For our paper we align the 5 Factors of Belonging in Mathematics with Teaching Mathematics with Social Justice (specifically the Teaching Practices), a broadened understanding of what it means to do mathematics (specifically the Learning Practices) as well as the Diversity Domain of the Social Justice Standards.

METHODS

In this study we share one set of lessons specifically developed to allow students in a first-grade classroom to experience a sense of belonging while learning about themselves and others in their classroom. The main goal of the lesson was to provide an opportunity for children examine how old the children in their first-grade classroom were and to use mathematics to understand this context. This is a complex question for first graders as several of them did not know how old they themselves were or how old their classmates were. The math goals related to: *Asking Questions*, as well as *Collecting, Organizing, Representing, Reading and Analysing Data* and Modelling *with Mathematics.* The Social Justice Standard addressed was:

> **Diversity 7** Students will develop language and knowledge to accurately and respectfully describe how people (including themselves) are both similar to and different from each other and others in their identity groups. (Chiariello et al., 2016)

Together the idea was to allow the students to see themselves and others in the class in the data being collected and analyzed and thus experience a sense of belonging in the class, learning about themselves and others, and a sense of doing mathematics.

We launched Lesson 1 by reading a few pages from the book *The Same But Different* (Potter, 2021) and engaging the children in discussing similarities and differences amongst themselves in the classroom. This allowed students to explore diversity in their classroom.

For the exploration part the teacher posed the question: How old are the children in this class? After a suggestion by a student the class divided itself up into "boys and girls." They counted 10 boys and 10 girls. Then they decided that this was not helpful to answer the question. They next suggested each child should say their age one-by-one. The teacher recorded the ages as the students shared them and then the class collectively counted that there were thirteen 6-year old students and seven 7-year old students in the class (that day). The teacher then represented this data with sticky notes in a bar graph. (see Figure 1). Discussions included what would happen to a child on/after their birthday (they would move from one column to the next).

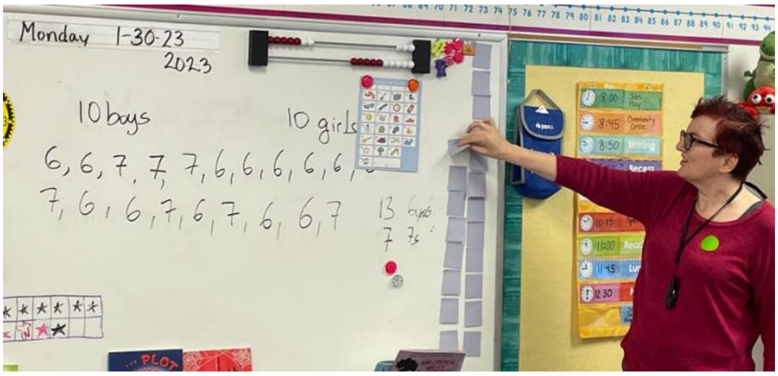


Figure 1: Image of data collected on student age both in numbers and a bar graph.

Lesson 2 launched with a Sesame Street video in which Elmo and Abby explore how they are the same and different [Sesame Street: Same and Different with Elmo and Abby](). Then the teacher presented the students with a bar graph made of photos of the students, so they could literally see themselves in the data (see Figure 2a) and examine the data. There were 22 students present that day. One student had had a birthday between the lessons so there was a discussion of how his image moved.

This activity led into the class creating four questions they were interested in and collecting, organization, and representing data (on posters) to answer those questions. Each poster had a heading decided by a small group of students who then labeled the x axis and put their own pictures on it. The questions examined were: *How many pets do you have? How many teeth have you lost? What is your favorite fruit?* and *What is your favorite Sonic Character?* (see Figure 2b). After each group initially created their poster with their question the students moved around the room to include their own picture on all the posters. Sometimes students fit into the predesigned categories; sometimes they needed to add additional categories. This resulted in 4 posters on which the students presented the answers to the question with their own data.

Figure 2a: Students verifying their own data is part of the data on the class.

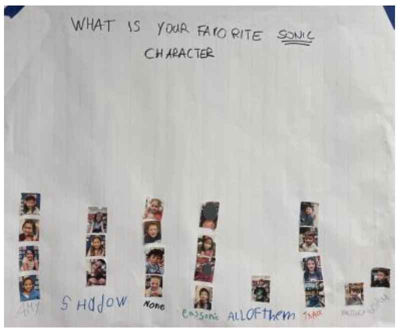


Figure 2b: Final poster on What is your favourite sonic character

Both lessons were videotaped, transcribed for analysis, and coded with MAXQDA software using the codes in Table 1. For an example of a coded transcript piece see Table 2.

Teaching Practices	**Student Learning Practices**	**Belonging**
maintain a focus on real-world reasoning	Students seeing/learning about themselves and others in the class	You see me
maintain a focus on mathematical reasoning	Students share ideas	I see myself
facilitate socio-political discourse		Others see me
facilitate mathematics discourse	**Social Justice Standards**	I am valued
elicit and use evidence of student thinking	SJS - Diversity 7	I can grow
	Mathematics Goals	
Asking Questions/Posing Problems	Collecting Data	Organizing Data
Representing Data	Reading Data	Analyzing Data

Table 1: Codes for MaxQDA coding.

Transcript	**Codes**
Leilani: because last time Benjamin wasn't seven and Benjamin is seven now. And the seven line was smaller because Benjamin wasn't there. Now Benjamin is seven. And Benjamin is in this line that means that is the seven line.	Math - Reading Data Learning Practice - Students share ideas. Learning Practice - Students seeing/learning about themselves and others in the class Belonging - Others see me
Leilani: And I am six and I do not see myself.	
Classmate: I see you Leilani you are right there	Belonging - Others see me Learning Practice - Students seeing/learning about themselves and others in the class
Leilani: Oh yeah	Belonging - I see myself

Classmate: And I am right here.	Belonging - I see myself

Table 2: Example of coding for a piece of transcript.

In the figures below we share images of the coded timeline for Lesson 1 (Figure 3), the visualization of the co-occurrence of the codes across Lesson 1 (Figure 4). We highlight the co-occurrence of the code **Students share ideas** with the **SJS - Diversity 7 code** as well as the **Belonging** codes (red circle). It also highlights the tight connection of the code **Students seeing/learning about themselves and others in the class** with the **Belonging** codes (green circle). In Figures 5 and 6 we focus the image of the coded timeline on the co-occurrence identified in Figure 4.

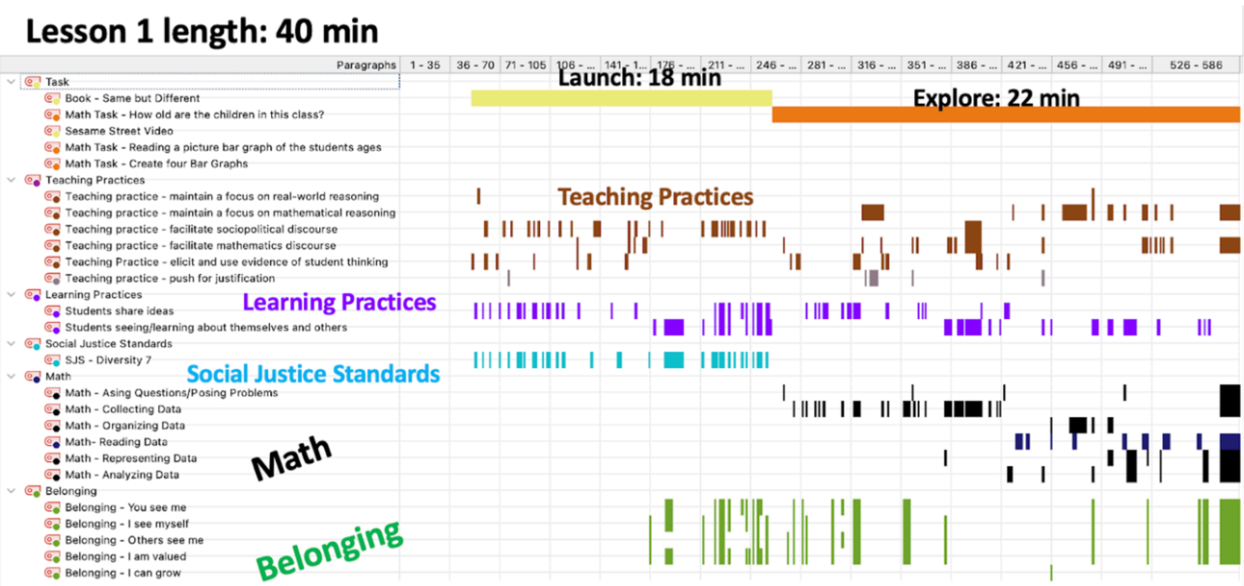


Figure 3: Timeline of Lesson 1 with coded segments

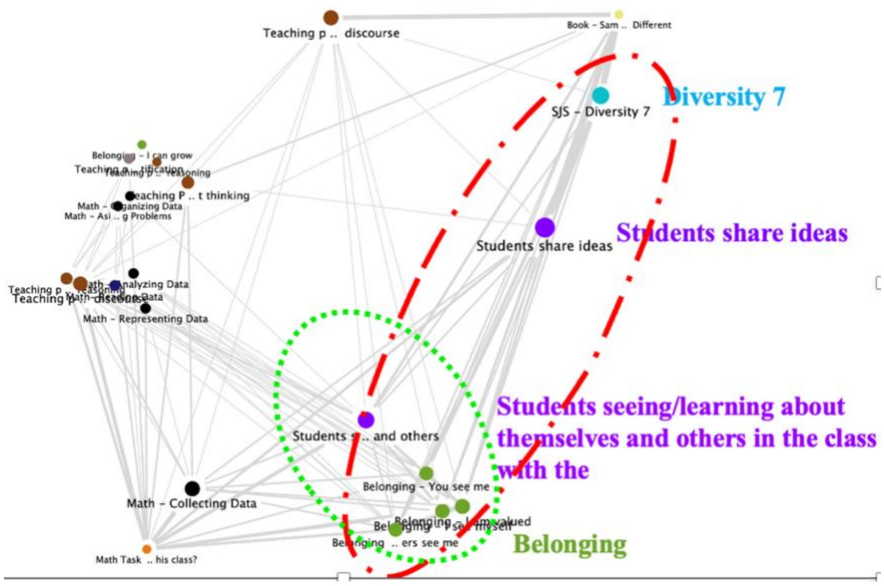


Figure 4: Co-occurrence of codes for Lesson 1. Each grey line represents a co-occurrence of codes. Red circle: Students share ideas with the **SJS - Diversity 7 code**

as well as the Belonging codes. Green circle: Students seeing/learning about themselves and others in the class with the Belonging codes.

This relationship can also be seen in Figures 5 and 6 which are selected codes in a timeline.

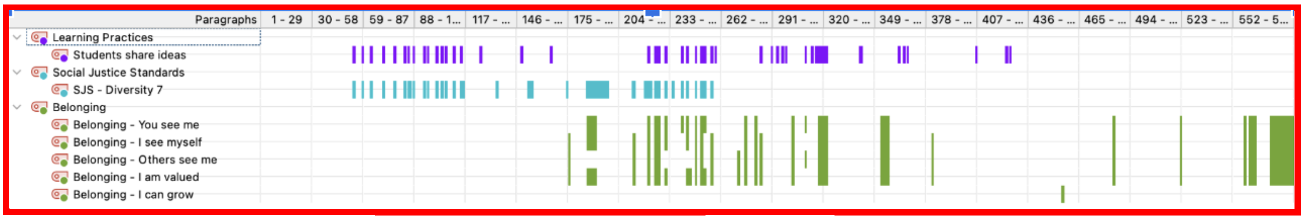


Figure 5: Co-occurrence **Students share ideas** with the **SJS - Diversity 7 code** as well as the **Belonging** codes (red circle in Figure 5)

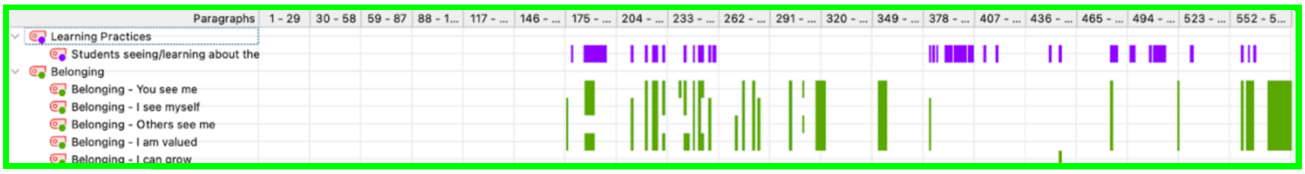


Figure 6: Co-occurrence **Students seeing/learning about themselves and others in the class** with the **Belonging** codes (green circle).

RESULTS AND INTERPRETATIONS

To answer our research questions: How can we create a sense of belonging for children in the elementary mathematics classroom? More specifically, can using data about themselves create a sense of belonging for the children in an elementary mathematics classroom? We analyzed both lessons with respect to all the codes. We highlighted in the methods above the co-occurrence of the learning practices codes **Students seeing/learning about themselves and others in the class** and **Students share ideas** with the code for **Belonging.** These codes co-occurred much of the time as seen in the example transcript, lesson visualizations and code relations above. Leilani (see transcript above) knew she belonged in the class and she was supposed to see herself; when she did not, she voiced her concerns. Her classmate helped her find herself in the data. Thus, creating tasks that allow students to see themselves (either in the data or literally in pictures within the data) can support students' sense of belonging. Further, students expect to see themselves in the data (because they *belong*).

The content of these lessons, while interesting to the students, was not critically focused for the most part. As such the **SJS - Diversity 7 code** was highlighted only during the launch activity of Lesson 1 (when discussing the book, *The Same but Different*) but not in the math activities across the rest of the two lessons. However, this book may lay a good foundation for examining more critical issues about themselves and others in the class and thus connecting the social justice standards throughout the lesson.

Across both lessons the math goals were consistently touched upon thus connecting mathematics learning to learning about themselves and their world. At the PME presentation we will include a more complete discussion of the data including the teaching practices. Lessons like these or the *Name Task* (Thanheiser et al., 2023) allow all students to experience a sense of belonging as they are designed to allow all students to see themselves in the data.

References

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

Bartell, T., Wager, A., Edwards, A., Battey, D., Foote, M., & Spencer, J. (2017). Toward a framework for research linking equitable teaching with the standards for mathematical practice. Journal for Research in Mathematics Education, 48(1), 7-21.

Chiariello, E., Olesen Edwards, J., Owen, N., Ronk, T., & Wicht, S. (2016). Social Justice Standards: The Teaching Tolerance Anti-Bias Framework. Teaching Tolerance: A Project of the Southern Poverty Law Center. Montgomery AL: Teaching Tolerance.org.

Conway IV, B. M., Id-Deen, L., Raygoza, M. C., Ruiz, A., Staley, J. W., & Thanheiser, E. (2022). Middle school mathematics lessons to explore, understand, and respond to social injustice. Corwin Press.

Cwik, S., & Singh, C. (2022). Students’ sense of belonging in introductory physics course for bioscience majors predicts their grade. Physical Review Physics Education Research, 18(1), 010139. https://doi.org/10.1103/PhysRevPhysEducRes.18.010139

Freire, P. (2018). Teachers as cultural workers: Letters to those who dare teach. Routledge.

Gopalan, M., & Brady, S. T. (2020). College students’ sense of belonging: A national perspective. Educational Researcher, 49(2), 134-137.

Potter, M (2021). The same but different (S. Jennings, Illus.). Featherstone.

Strayhorn, T. L. (2018). College students' sense of belonging: A key to educational success for all students. Routledge.

Strayhorn, T. L. (2019). College students' sense of belonging: A key to educational success for all students (2nd ed.). Routledge.

Thanheiser, E. (2023). What is the Mathematics in Mathematics Education? The Journal of Mathematical Behavior, 70, 101033.

Thanheiser, E., Koestler, C., Sugimoto, A., & Felton-Koestler, M. (2023). What’s in a name? Collecting, organizing, representing data. Mathematics Teacher: Learning and Teaching Pre-K–12. 116(10), 746-752.

Trujillo, G., & Tanner, K. D. (2014). Considering the role of affect in learning: Monitoring students' self-efficacy, sense of belonging, and science identity. CBE—Life Sciences Education, 13(1), 6-15.

Valero, P (2023). Mathematical subjectivation: Death sentence of chances for a terrestrial life? Ayalon, M.; Koichu, B.; Leikin, R.; Rubel, L. & Tabach, M. (Eds.) Proceeding of the 46th Conference of the International Group for the Psychology of Mathematics Education (Vol. 1). University of Haifa, Israel: PME.

ANALYSIS OF THE COGNITIVE ACTIVATION OF COMBINATORIAL TEXTBOOK TASKS IN GRADE 11 AND 12

Charlott Thomas and Birte Pöhler

University of Potsdam

Various studies have shown that textbooks and their tasks are essential for mathematics learning. This also applies to combinatorics, with which learners often have difficulties. Accordingly, the combinatorics chapters of five textbooks for the upper secondary school level in Germany were analyzed. The analysis revealed that the textbook tasks predominantly require procedural thinking and show little variety in other task types (declarative and conceptual thinking). The homogeneity of combinatorics tasks in textbooks must be revised as it may impair students' cognitive activation, an essential aspect of teaching quality.

INTRODUCTION

The cognitive activation of students, one of the three central characteristics of teaching quality, is considered a decisive factor for students' learning gains (Fauth & Leuders, 2022). However, it is problematic that the tasks used in mathematics classrooms often have a low overall cognitive activation potential (Jordan et al., 2008). With the help of the system of categories on the cognitive activation potential of mathematical tasks by Neubrand et al. (2013), the aim is to investigate whether this also applies specifically to combinatorics. This could provide an initial indication of why students have problems solving combinatorial tasks.

THEORETICAL BACKGROUND

Students have significant challenges by solving combinatorial problems (e.g., Annin & Lai, 2010). However, textbooks usually classify counting problems according to the various basic combinatorial situations (permutation, variation, and combination, each with and without repetition) (Fischbein & Gazit, 1988; Batanero et al., 1992) in order to provide learners with clear guidelines and procedures. Since the structure and the content of the existing textbook guides most teachers or they use it primarily as a collection of tasks (e.g., Haggarty & Pepin, 2002), combinatorial tasks from textbooks commonly used in Germany will be analyzed regarding their cognitive activation. Learning opportunities are described as "cognitively activating" if they encourage learners to engage actively with the learning content, which happens at an optimal level for them (Krauss et al., 2004). Therefore, Cognitive activation is a decisive factor in students' learning gains (Fauth & Leuders, 2022). For teaching, this means that students' different cognitive prerequisites should be taken into account, the learning time should be used optimally for the anticipated competence development, and the students should be encouraged to engage in challenging cognitive activities that focus on the mathematical content (Leuders & Holzäpfel, 2011). It is problematic that past

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. 4, pp. 153–160). PME.

studies have found a low cognitive activation potential of tasks in German mathematics classrooms (Jordan et al., 2008, p. 99).

In order to provide an appropriate assessment of the level of difficulty and the cognitive activation potential of tasks, a specific system of categories for mathematical tasks was developed as part of the project "Teachers' professional knowledge, cognitively activating mathematics teaching and the development of mathematical competence" (COACTIV) (Jordan et al., 2006; Neubrand et al., 2013). Originally, this distinguishes between four dimensions. As the subject area of the tasks to be taken into account in this article is fixed, Dimension A – Content Framework is not considered. On the one hand, this describes the material extent of cognitive activation about the individual subject areas (topic areas) and, on the other hand, the extent to which tasks from previous grade levels (curricular knowledge) are included. Accordingly, the focus of this article is on the dimensions B – Cognitive Framework, C – Elements of the modeling cycle, and D – Search space for the solution, as these provide direct insight into the cognitive activation potential of tasks, particularly in connection with problem solving and modeling (e.g., Blum & Leiß, 2007).

In order to map the different requirements that can be contained in mathematical tasks, the solution process was categorized into three *Types of mathematical activities* in *Dimension B – Cognitive Framework*: *technical task* (demand for skills or declarative knowledge), *procedural task* (predominantly procedural thinking required) and *conceptual task* (predominantly conceptual thinking required) (e.g., Neubrand, 2002). These three types form the core of the competency model used to evaluate the OECD PISA test in Germany (Neubrand, 2002). It should be noted that there can be both easy and challenging tasks in all three task types.

The theoretical model of the modelling cycle (Blum et al., 2007) forms the basis for the differentiated representation of cognitive activities in *Dimension C – Elements of the modelling cycle*. *Extra-mathematical modelling* addresses translation processes between reality and mathematics, i.e., the mathematization of real-life situations and the interpretation/validation of mathematical results. This extra-mathematical modelling can take place directly by specifying a model *(standard modelling)*, not directly and when mathematizing with several steps *(multistep modelling)* as well as through the validation, reflection and/or assessment of mathematical models *(reflection on a model, development, and validation of complex models)*. Within mathematics, translation processes can also take place *(inner-mathematical modelling)*. For example, only one specific object or one modelling step may be necessary to solve the task *(standard modelling)*. Further knowledge from other mathematical (sub-)fields may be required, or several modelling steps may be carried out *(multistep modelling)*. Furthermore, it can be encouraged to design a comprehensive strategy so that general statements are made, or solutions are critically reflected upon *(reflection on a model, validation, and strategy development)*. The importance of the complexity of the task text has been proven to be a difficulty-generating feature of PISA tasks (Cohors-Fresenborg et al., 2004). When dealing with mathematical texts, it is, therefore, essential to consider the extent to which the linguistic sequence of the task corresponds to the solution path in a mathematical model. In the *Processing of mathematical texts*, for example, the primary and subordinate clauses may correspond to the steps of mathematical processing and linguistic references may be made *(direct text comprehension)*, the order of the sentences may correspond to the steps of mathematical processing in a more challenging way or not at all *(text comprehension with reorganization)* or logical functions and sophisticated linguistic techniques may be necessary for understanding the text *(comprehension of logically complex texts)*. Argumentation plays a role in both intra- and extra-mathematical processes, whereby a distinction is made between the mere reproduction of standard argumentation *(standard reasoning)*, manageable multistep argumentation *(multistep argumentation)*, or complex mathematical argumentation and its comparison and reflection *(development of complex argumentation, proofs, evaluation of argumentations)* (Jordan et al., 2006).

In *Dimension D – Search space for solution*, the *Direction of task solution* of the tasks and the *Number of solution paths required* are recorded. An important feature here is the extent of the solution space, which is based on two criteria: First, whether a task is set according to or contrary to the usual direction of learning or representation of a concept or procedure *(forward* or *backward)*. Secondly, whether a task requires different solution paths *(none, one,* or *several solution paths)*. Tasks that fulfill the last aspect play a significant role in cognitive activation in the classroom (Blum & Wiegand, 2000).

RESEARCH QUESTION AND METHODICAL APPROACH

German textbooks for grades 11 and 12 contain many combinatorial tasks so that students can be given sufficient learning opportunities to deal with these in mathematics classroom. At the same time, various studies show that students face significant challenges when they have to solve combinatorial problems (Annin & Lai, 2010). Against this background, this article aims to investigate the extent to which the cognitive activation potential of the tasks to be solved in the textbooks could be responsible for these difficulties. Specifically, the following research question will be addressed:

> *To what extent do 11th and 12th grade combinatorial textbook tasks prove to be cognitively activating?*

A total of five textbooks from two publishers that are among the largest in Germany were examined, namely Mathematics 11 Basic Course, Mathematics 11 Basic Course and Advanced Course (state-specific) and Mathematics 12 Basic Course and Advanced Course (state-specific) as well as the Lambacher Schweizer Gesamtband (nationwide).

A total of two textbooks were coded by two independent coders. The intercoder reliability rate was determined using MAXQDA and amounted to .77 and .84, respectively, which is why a predominant consensus in the data analysis can be assumed. These values are considered very good according to conventional standards (Döring & Bortz, 2016).

The tasks in the selected 11th and 12th grade textbooks in Germany were analyzed based on dimensions B to D of the system of categories by Neubrand et al. (2013) presented above. Based on this, the system of categories was adapted and used as an evaluation tool for the textbook analysis:

Dimension	Category	Properties
B – Cognitive Framework	Type of mathematical activity	1 = technical task, 2 = procedural task, 3 = conceptual task
C – Elements of the Modelling cycle	Extra-mathematical modelling	0 = not required, 1 = standard modelling, 2 = multistep modelling, 3 = reflection on a model, development and validation of complex models
	Inner-mathematical modelling	0 = not required, 1 = standard modelling, 2 = multistep modelling, 3 = reflection on a model, validation, strategy development
	Processing of mathematical texts	0 = not required, 1 = direct text comprehension, 2 = text comprehension with reorganization, 3 = comprehension of logically complex texts
	Argumentation	0 = not required, 1 = standard reasoning, 2 = multistep argumentation, 3 = development of complex argumentation, proofs, evaluation of argumentations
D – Search space for the solution	Direction of task solution	1 = forward, 2 = backward (“reverse task”)
	Number of solution paths required	0 = none, 1 = one, 2 = several

Table 1: Overview of selected categories of the system of categories, including the properties (adapted from Neubrand et al., 2013)

Because many tasks consisted of questions with subtasks, subtasks were treated as independent units of analysis. Accordingly, subtasks of a task are considered as independent tasks if they are numbered and have their own instructions. On the other hand, tasks that contain several sub-instructions or work steps are regarded as one task. This definition of the units of analysis corresponds to that used in COACTIV (Jordan et al., 2006). Only tasks for which there are no solutions are analyzed, therefore example or sample tasks are not examined.

RESULTS

When analyzing the five textbooks, it was found that both the textbooks Mathematics 11 Advanced and Basic Course and Mathematics 12 Advanced and Basic Course list the same tasks in the combinatorics chapter. These textbooks are therefore listed together below. In total, all 108 combinatorial tasks in the textbooks (43 tasks in Mathematics 11, 34 tasks in Mathematics 12, and 33 tasks in the Lambacher Schweizer) were analyzed about the categories according to Neubrand et al. (2013).

The combinatorial tasks in the textbooks are very similar in their classification in the system of categories of the COACTIV study. In the following, the coding for the task below (Figure 1) will be carried out as an example.

> Photo
>
> How many ways are there to arrange eleven players of a soccer team in a line for a photo?

Figure 1: Combinatorial example task from Mathematics 12, translated by the author

This task is a *procedural task* because it requires modelling, and the processing phase mainly requires procedural thinking. Furthermore, the extra-mathematical and inner-mathematical translations prove to be *standard modelling*. The extra-mathematical translation can be carried out directly, as the model is close at hand, and the inner-mathematical translation can solve the task with one modelling step, as only one concrete object is considered. In addition, *direct text comprehension* is assumed here, as the existing main and subordinate clauses (in German) are arranged in the order of mathematical processing, which is also reflected in the direction of task solution as the discussion proceeds along the usual line of thought *(forward)*. Furthermore, on the one hand, no argumentation is required, as no types of reasoning are required and, on the other hand, no explicit solution paths are required *(none)*.

The analysis of all tasks shows the following results:

Category in **Dimension B**	Mathematics 11	Mathematics 12	Lambacher Schweizer
Type of mathematical activity			
1 = technical task	9%	0%	3%
2 = procedural task	88%	97%	91%
3 = conceptual task	2%	3%	6%

Table 2: Analysis of the textbook tasks in Dimension B

About *Dimension B*, the category *Type of mathematical activity* shows that the tasks in the textbooks are mainly *procedural tasks* that require modelling and procedural thinking. On the other hand, there are only a few tasks in which conceptual thinking (conceptual task) in the processing phase or in which only technical knowledge (technical task) is required.

Dimension C in the category *Extra-mathematical modelling* mainly contains tasks that require *standard modelling* (63% in Mathematics 11, 68% in Mathematics 12, and 73% in Lambacher Schweizer). This means that the tasks mainly require direct translations between the actual situation and the mathematical world, as the model is (directly) given. Only about a third of the tasks in the respective books require *multistep modelling* (28% in Mathematics 11, 29% in Mathematics 12, and 24% in the Lambacher Schweizer), where mathematization involves several steps or different mathematical topics. In the *Inner-mathematical modelling* category, the tasks almost exclusively require standard modelling (86% in Mathematics 11, 97% in Mathematics 12, and 94% in Lambacher Schweizer), where the approach is implicitly suggested in the task and only one modelling step is required to solve it. In the Mathematics 11 textbook, no inner-mathematical modelling is required in 9% of the tasks; these tasks prove to be technical tasks in which only technical knowledge is required without any contextual connection. In the category *Processing of mathematical texts*, around half of the tasks (58% in Mathematics 11, 44% in Mathematics 12, and 52% in Lambacher Schweizer) are characterized by few or hardly any text. In these tasks, the order of the sentences corresponds to the steps of the mathematical processing, and the task texts consist mainly of simple main sentences without subordinate clauses. The other half of the tasks (42% in Mathematics 11, 56% in Mathematics 12, and 48% in the Lambacher Schweizer) are characterized by *direct text comprehension*, in which the order of the sentences does not directly correspond to the steps of the mathematical processing and the task consists of several main and subordinate clauses, so that linguistic references must also be made. In the *Argumentation* category, almost all tasks analyzed can be characterized by the fact that no argumentation is necessary (98% in Mathematics 11, 97% in Mathematics 12, and 97% in Lambacher Schweizer). There is only one task per textbook in which *standard reasoning* is required. In this low-level reasoning, standard reasoning is reproduced, e.g., the reasoning must be developed in one step or purely mathematically.

Category in **Dimension D**	Mathematics 11	Mathematics 12	Lambacher Schweizer
Direction of task solution			
1 = forward	100%	100%	100%
2 = backward ("reverse task")	0%	0%	0%
Number of solution paths required			
0 = none	100%	100%	100%
1 = one	0%	0%	0%
2 = several	0%	0%	0%

Table 2: Analysis of the textbook tasks in dimension C

In *Dimension D*, in the categories *Direction of task solution* and *Number of solution paths required*, the tasks are predominantly designed in such a way that, on the one hand, calculations should conform to the common direction of mathematical concept. On the other hand, no concrete or alternative solution path is required for any task (100% in all textbooks). For the former, this means that the discussion of the tasks follows the usual direction of thought in mathematics and that they cannot be described as reversal tasks. The second means that the textbooks neither specify a concrete path in the tasks nor require several possible solutions.

DISCUSSION AND CONCLUSION

Mathematical content that is taught – such as combinatorics – is often experienced through the working on tasks, and the students' mathematical activity is usually concentrated on dealing with these tasks. They are a vehicle for the students' cognitive activities (Jordan, 2006). In the textbook chapters on combinatorics, students can also be confronted with various tasks that provide the basis of potential learning opportunities. The analysis of the selected popular textbooks for the upper secondary level in Germany has shown that all five textbooks offer students very similar tasks.

Overall, it is also clear that the tasks in the chosen textbooks are exceptionally homogeneous and not very varied. There are no significant differences between the state-specific and the nationwide textbooks. Although only a selection of textbooks was examined, the analyses provide initial indications that the tasks contained in the textbooks could be more cognitively activating. Since the cognitive activation of students is seen as a decisive factor for their learning gains, this monotony of tasks needs to be revised. Since most teachers use the textbook as a collection of tasks and observation studies show that the lessons need to be cognitively activating (OECD, 2020), this study can provide initial indications that the lessons are not as cognitively activating, even if only selected textbooks were analyzed. Teachers should be aware of this problem and provide learners with further learning opportunities so that they can handle the challenges. The balanced complexity of the individual dimensions (e.g., through a quantitatively balanced range of tasks) is essential to activate students cognitively.

References

Annin, S., & Lai, K. (2010). Common Errors in Counting Problems. *The Mathematics Teacher*, 103(6), 403–409.

Batanero, C., Navarro-Pelayo, V., & Godino, J. D. (1997). Effect of the implicit combinatorial model on combinatorial reasoning in secondary school pupils. *Educational Studies in Mathematics*, 32(2), 181–199.

Blum, W., & Wiegand, B. (2000). Offene Aufgaben – wie und wozu? In *mathematik lehren*, 100, 52–55.

Chi, M. T. H. & Wylie, R. (2014). The ICAP Framework: Linking cognitive engagement to active learning outcomes. *Educational psychologist*, 49(4), 219-243.

Cohors-Fresenborg, E., Sjuts, J., & Sommer, N. (2004). Komplexität von Denkvorgängen und Formalisierung von Wissen. In M. Neubrand (Ed.), *Mathematische Kompetenzen von Schülerinnen und Schülern in Deutschland. Vertiefende Analysen im Rahmen von PISA 2000* (1. Auflage, pp. 109-144). VS Verlag für Sozialwissenschaften.

Döring, N., & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften.* Springer Verlag.

Fauth, B., & Leuders, T. (2022). *Kognitive Aktivierung im Unterricht*. Institut für Bildungsanalysen Baden-Württemberg.

Fishbein, E. & Gazit, A. (1988). The combinatorial solving capacity in children and adolescents. *Zentralblatt für Didaktitk der Mathematik* 5, 193-198.

Haggarty, L. & Pepin, B. (2002). An Investigation of Mathematics Textbooks and their Use in English, French and German Classrooms: Who gets an opportunity to learn what?. *British Educational Research Journal*, 28(4), 567-590.

Jordan, A., Ross, N., Krauss, S., Baumert, J., Blum, W., Neubrand, M., Löwen, K., Brunner, M., & Kunter, M. (2006). *Klassifikationsschema für Mathematikaufgaben: Dokumentation der Aufgabenkategorisierung im COACTIV-Projekt.* Max-Planck-Institut für Bildungsforschung.

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

Krauss, S., Kunter, M., Brunner, M., Baumert, J., Blum, W., Neubrand, M., Jordan, A., & Löwen, K. (2004). COACTIV: Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung von mathematischer Kompetenz. In *Bildungsqualität von Schule: Lehrerprofessionalisierung, Unterrichtsentwicklung und Schülerforderung als Strategien der Qualitätsverbesserung* (pp. 31-53). Waxmann.

Leuders, T., & Holzäpfel, L. (2011): Kognitive Aktivierung im Mathematikunterricht. *Unterrichtswissenschaft,* 39(3), 213-230.

Neubrand, J. (2002). *Eine Klassifikation mathematischer Aufgaben zur Analyse von Unterrichtssituationen. Selbsttätiges Arbeiten in Schülerarbeitsphasen in den Stunden der TIMSS-Video-Studie.* Franzbecker.

Neubrand, M., Jordan, A., Krauss, S., Blum, W., & Löwen, K. (2013). Task Analysis in COACTIV: Examining the Potential for Cognitive Activation in German Mathematics Classrooms. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Eds.), *Mathematics teacher education: Vol. 8. Cognitive activation in the mathematics classroom and professional competence of teachers: Results from the COACTIV project* (pp. 125–144). Springer.

OECD. (2020). *Global Teaching InSights A Video Study of Teaching*. OECD Publishing.

Valverde, G. A. (2002). *According to the book: Using TIMSS to investigate the translation of policy into practice through the world of textbooks*. Kluwer Academic Publishers.

LANGUAGE AS A TRANSPARENT RESOURCE FOR DEVELOPING MATHEMATICAL UNDERSTANDING

Pauline Tiong

Simon Fraser University

While the notion of language as a resource is not new and of increasing interest in mathematics education research, not many researchers focus on understanding the notion from the perspectives of teachers. Motivated by an interest to understand the existing state of how teachers are noticing and using language (particularly the mathematics register) as a resource in the mathematics classroom, this paper reports the findings from a task-based interview conducted with one teacher. By accounting for the teacher's responses to the interview through the lens of language-related dilemmas and orientations, I was able to glean insights into how she notices and uses language (particularly the mathematics register) as a transparent resource primarily for developing understanding in the teaching and learning of mathematics.

LANGUAGE AS A RESOURCE IN THE MATHEMATICS CLASSROOM

Since the 1990s, the notion of *language as a resource* for the teaching and learning of mathematics has been crucial in shifting the deficit-oriented perspective to language in mathematics education research, particularly in the context of multilingual classrooms (e.g., Adler, 2002; Moschkovich, 2002). Most researchers have also moved beyond asking the question of whether language can be a resource in mathematics education. Instead, the question they are asking in recent research is how better to help teachers use language as a resource by focusing on the development of teachers' language-responsiveness in mathematics teaching and learning (e.g., Adler, 2021; Prediger et al., 2019). Considering Vygotsky's (1934/1986) works which have elaborated at length the intricate connections between thought *(mathematical thinking)* and *(mathematical)* language, it is indeed necessary for teachers to be cognizant of how to use language better as a resource in the mathematics classroom.

However, there seems to be little focus in understanding the existing state of how teachers are noticing and using language as a resource in their classrooms. Moreover, *language* in most research has become more encompassing and may not only attend to the *mathematics register* (Halliday, 1975; Pimm, 1987), which more accurately represents what is commonly referred to as mathematical language. Notably, the *mathematics register* is deemed to serve the function of thinking about (and communicating in spoken or written forms) mathematical ideas and meanings (Pimm, 1987). Thus, the mathematics register can be considered a tool for thinking and communicating mathematics (Vygotsky, 1934/1986), and viewed as an important resource for mathematics education. This led to my attempt in wanting to study how teachers are noticing and using language (particularly the mathematics register) as a

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. 4, pp. 161–168). PME.

resource for mathematics teaching and learning. Specifically, I seek to explore this phenomenon through the lens of teachers' language-related dilemmas and orientations.

TEACHERS' LANGUAGE-RELATED DILEMMAS AND ORIENTATIONS

Language-related teaching issues or dilemmas were first proposed by Jill Adler (2002) in her work, primarily in the contexts of multilingual classrooms. The notion of teaching dilemmas, first proposed by Lampert (1985), refers to situations of tensions which teachers may face in their teaching practice when there seems to be "no one 'right' solution" (Adler, 2002, p. 49) that can resolve the tensions, from the perspective of teachers. Building on this, Adler identified three dilemmas teachers face in relation to the use of language in (multilingual) mathematics classrooms, namely: (a) the *dilemma of code-switching* where teachers need to decide whether to change the language of instruction to develop students' mathematical understanding when decisions made to change the language of instruction may compromise the learning of the mathematics register; (b) the *dilemma of mediation* where teachers need to decide whether to intervene to validate students' meanings during group discussions or presentations when decisions made to intervene may compromise students' opportunities to develop mathematical communicative competence; (c) the *dilemma of transpar*ency where teachers need to decide whether to teach the mathematics register when decisions made to teach the language explicitly (or making language a visible resource) may compromise the development of student mathematical understanding (where language works as an invisible resource). While these dilemmas were surfaced through Adler's work in multilingual classrooms, she suggested that they can be similarly faced by any teacher who uses the mathematics register as a resource for teaching and learning. Correspondingly, Zazkis (2000) has extended the dilemma of code-switching to include situations when teachers may need to decide between the use of the mathematics register and everyday language, rather than across different languages in a relatively monolingual context.

By contrast, Susanne Prediger et al. (2019) discussed how teachers' orientations can often influence their practices. In other words, a teacher's language-related orientations are likely to lead to a different focus or treatment, in terms of pedagogical approaches and actions, of language as a resource for mathematics teaching. Consequently, they identified five language-related orientations as being crucial in influencing teachers' practices in the mathematics classroom. These include the extent which mathematics teachers assume responsibility for *language learning as a goal*; *strive for pushing rather than reducing language* in relation to language demands (such as noticing, supporting and developing language in the context of mathematics teaching); *focus on the discourse level rather than on word level only* in learning the mathematical language; have *integrative perspectives instead of additives only* to the learning of language in teaching mathematics; focus on *conceptual understanding before procedures* which necessitates the use of language (i.e., the mathematics register) as a resource for mathematics teaching.

While both theoretical constructs – teachers' language-related dilemmas and language-related orientations – have been used to study teachers' use of language in mathematics classrooms, they each focus on a very specific aspect that inform how teachers notice and use of language as a resource in mathematics teaching and learning. As such, I see value in how both constructs can be used in a complementary manner to account for what and how teachers notice and use language (particularly the mathematics register) as a resource in the teaching and learning of mathematics.

A TASK-BASED INTERVIEW WITH KAREN

Data for this paper was drawn from my dissertation study, which focused on addressing the phenomenon of interest using task-based interviews. A total of eleven experienced mathematics teachers were interviewed and their teaching contexts ranged from elementary school level to tertiary level. Each task-based interview included a series of reflection tasks (Zazkis & Hazzan, 1998) designed to illuminate specific situations for teachers to reflect on their own practice and experiences (Mason, 2002), in terms of how language would be used in their mathematics classrooms. The tasks were presented in the form of dialogues which simulated situations with language-related dilemmas in relation to the use of the mathematics register (Adler, 2002). In particular, the teachers were asked to reflect upon what they noticed, in terms of the language students were using, and how they would respond if they were the teacher in the situations illustrated. To analyse the interview data, I adopted Mason's approach of *account-of* and *accounting-for*, as the method to understand how the teachers would notice and use language in their mathematics classrooms. Based on the interview responses, I first created an *account-of* what each teacher noticed in terms of language use (particularly the mathematics register) and their corresponding actions/ reactions in relation to the given tasks. Subsequently, the *accounts-of* were analysed with the intent of *accounting-for* how and why they would use language in their classrooms through understanding their experiences with language-related dilemmas (Adler, 2002) and their language-related orientations (Prediger et al., 2019) respectively.

In my dissertation study, I observed two main categories in which language has typically been noticed and used by teachers in the mathematics classrooms, namely as a resource for developing mathematical understanding, and as a resource for mathematics talk. However, due to the size of this paper, I can only share a snapshot of the findings to the first category through the case of Karen, a university instructor who primarily considers language as a resource to develop mathematical understanding. A snippet from the account-of Karen's interview responses to one particular task is also presented to illustrate some of the findings.

Account-of Karen's responses to one task (graphs of rational functions)

The task depicts an (fictional) account of two secondary students, Ethel and Theo, discussing their observations about the dotted lines (asymptotes) in the graphs using (more) everyday language. As they were trying to recall the term *asymptote*, the teacher, Ms. Wilson, intervened and highlighted a series of register terms (in bold).

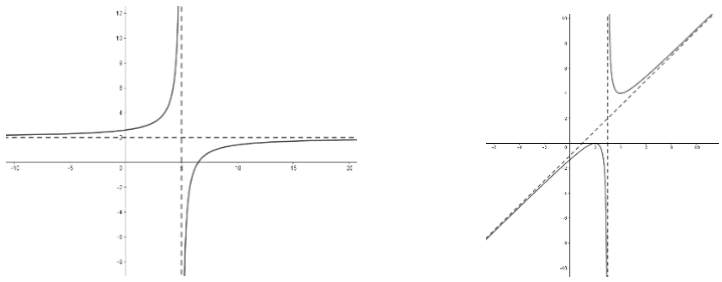


Table 1: Diagram of graphs given in the task.

Ethel: Look at these graphs, they all have these lines [pointing to the dotted lines] which the graphs go very near to. They can be straight up or lying down?

Theo: I see them too. I think Ms. Wilson used a special term. Eek, but I can't remember. But, are these the same thing? There's like more of them, there's a slanted one, and they can cut each other. Ms. Wilson, we have a question. These lines that the graphs go closer and closer to but can never touch or meet the lines. We know there's a name for it but ....

Ms Wilson: Ah, you are talking about asymptotes, **A-SYMP-TOTES**. Yes, these lines are all asymptotes. Both of you made the good observation that the graphs are **A-PPROA-CHING** these lines, asymptotes, going closer and closer without touching, or rather, **IN-TER-SEC-TING** them. They can be **VER-TI-CAL**, what you mean by straight up, **HO-RI-ZON-TAL**, what you mean by lying down and also **OB-LIQUE**, for those slanted lines.

In the interview, Karen first noticed how Ethel and Theo were "noticing the difference between the dotted lines and the other lines" and listing the various properties of the dotted lines (asymptotes), based on their reading of the graphs. To her, the students seemed to know "what's happening, they just don't have the word for it, they're actually reaching for the word" or the "classification" for the dotted lines in the task. She immediately shared that, unlike the other tasks, it would be a "more clear-cut" situation where "terms in math show up". Hence, she would step in and say, "that's an asymptote", as it would be "helpful to talk about it". Other than introducing or reminding students of the term *asymptote* they seemed to be looking for, she added that she would ask students to define *asymptotes* and clarify if the properties they had listed are true for all asymptotes. Upon reading what Ms. Wilson said in the task, Karen first expressed hesitance in doing what Ms. Wilson did, which was to correct "things they (students) have said that aren't wrong, just maybe aren't as precise as they could be". Karen added that she might even use terms such as "these guys" to refer to the dotted lines prior to introducing the term *asymptotes* to "make math more approachable to students". Depending on her rapport with the students, she would probably introduce mathematics terms gradually (instead of at once) while code-switching with what the students said and emphasising that what they said were not wrong. She added that "tone matters quite a bit" when introducing new or proper terms to students. Other than not correcting students at once, she would instead use gesturing or questioning to better understand what students were thinking first, before introducing the terms.

She further elaborated on her reluctance to "gate keep" the use of language or be "privileging certain terms over another without any obvious benefit". She explained, "words are right because we've decided they're right, as mathematicians". If what the students said meant the same idea, she would lean towards a more "casual" way of communicating in her classroom, while introducing formal language when necessary. For example, she would introduce terms either when students asked for them (e.g., asymptote) or when the terms help to define or make ideas clearer or simpler (e.g., approaching, vertical, horizontal) for presentations or written work. To her, "that's how terms in math show up" and they are not "defined for fun". As Karen mentioned the use of more precise use of language during students' presentation and written work, as compared to students' discussions, I asked if that was her preference to how students use language in her classroom. She commented that her preference would depend on "how formal (oral and written) it is and who your audience is", elaborated that she would often tell students that there are three "different levels of convincing people", which include themselves, a friend and an enemy and explained that language use would need to be increasingly precise from everyday or colloquial to more formal or mathematical when convincing oneself to a friend to an enemy. She gave an example on how it would be "too much" to "throw in nine different terms into a paragraph that is meant for someone who's not in the mathematical community".

When asked if there were instances when she had taught formal mathematical terms explicitly (like what Ms. Wilson did), Karen shared what she did when teaching differential equations before. At the start of the class, she told the students that she had to teach them a few terms, including *general solution*, *particular solution*, and *arbitrary constant*. Her reason (to her students too) was that these terms would be frequently used in subsequent lessons, and she could not be "saying the solution that comes up when you're solving a differential equation, and it has a constant in it" whenever she referred to the *general solution*. She explained that she would always "preface" the teaching of the terms or words in her class as "words sometimes exist without reasons" in the mathematics classroom and students might not know why they needed to know the words. She further mentioned that if she had to be "pedantic about it […] there's a mathematical community where students should know the words that other people are using". She added that it is her responsibility to help students "feel comfortable in the mathematical community", rather than "feel alienated by words they cannot use" or not know when they are learning mathematics.

Accounting-for Karen's case through language-related dilemmas

From the account-of Karen's responses to the various tasks, she did not seem to face any obvious tension though she noticed the respective language-related dilemmas (Adler, 2002) in the tasks. Except for one instance (near the end of the interview) when she began to question her stand in deciding not to mediate students' use of language when they were not wrong, she was generally clear in her considerations with regard to when to code-switch, mediate use of student language or teach language explicitly (particularly the mathematics register) in relation to the tasks.

In relation to the *dilemma of code-switching*, it is evident that she would use both the mathematics register and everyday language, and would not hesitate to code-switch between registers whenever necessary, in the teaching and learning of mathematics. Specific to how language should be used, she often stressed the importance of meeting students at their comfort level and readiness in using language to learn mathematics. For instance, she might even use “these guys” before introducing the term *asymptote*. She commented that introducing or new or difficult concepts using everyday or informal language first would help to “make math more approachable to students”.

In relation to the *dilemma of mediation,* Karen mostly demonstrated clarity in deciding when and how much she would mediate students’ use of language in relation to the tasks. Broadly, her key considerations would be students’ level of understanding and their need for language in communicating mathematical ideas. She would typically mediate the use of language in situations which she deemed as necessary to clarify students’ confusion or disagreements and deepen their understanding of the concepts to be learnt. She would also be more inclined to introduce or use the mathematics register if students had reached a certain level of understanding and required certain language to progress further in their discussion. For example, she would certainly provide students with the term *asymptotes*, because the students had showed awareness of the concept and had requested for the term in the task. She would then make use of the opportunity to (re-)introduce the term and re-affirm or further their understanding of *asymptotes.* But unlike Ms. Wilson, Karen would be hesitant to “correct” the other terms when the students were not wrong in their thinking and merely imprecise in their language. She would only introduce selective terms, such as *vertical* and *horizontal*, but not *oblique*. She explained that *vertical* and *horizontal* are simpler words, commonly used in everyday contexts, while there did not seem to be more benefits in using *oblique* instead of slanted. She would also introduce and use *approaching* which “means that things are getting closer together in a really simple way”. However, she emphasised that, “there’s nothing wrong with saying getting closer and closer without touching” and, hence, she would not insist students to use *approaching*, as it would be “privileging certain terms over another without any obvious benefit”. Thus, she would not be overly concerned if students did not use the proper language in their verbal discussions if they could clearly describe the mathematical concepts in their own words.

In relation to the *dilemma of transparency,* Karen was similarly decisive and articulate in responding to when and why she would teach the mathematics register explicitly. Again, if a decision to teach language explicitly were to value-add (and not compromise) the development of mathematical understanding and the ease of mathematical communication, she would generally be more inclined to do so. For example, she did not agree with Ms. Wilson’s approach of teaching language without involving students in the sense-making process of the terms. She opined that such an approach would be overwhelming for students and not meaningful pedagogically in helping students develop mathematical understanding. Nonetheless, there had been instances when she would “teach the mathematics register more” explicitly (make language visible) first. For example, when teaching differential equations, Karen shared how she taught the necessary terms, such as *particular solution* and *general solution*, to students right from the start and made it a point to explain why she needed to do so. In particular, (re-)explaining what the term *general solution* meant each time she had to refer to it would probably not be effective in the teaching and learning process. However, she added that her primary intent would be placed on building students' awareness of the register and help them be familiar and comfortable with the language of the mathematics community.

Accounting-for Karen's case through language-related orientations

On the whole, Karen's actions in managing the different language-related dilemmas collectively substantiated a positive inclination towards the five language-related orientations (Prediger et al., 2019). Firstly, *language is likely a learning goal* in Karen's classroom, based on how she would assume the responsibility to help students develop both understanding of mathematical concepts and the corresponding register required to talk about the concepts. This was evident in her example of the differential equations class. Moreover, her preference for students to be using language more precisely in formal settings and with a larger audience suggests that her students would likely be learning language alongside learning mathematics, in preparation for those situations. Thus, she would *unlikely reduce* language in her mathematics classroom. However, she emphasised that she would *not blindly push* for the use of certain language if there were to be no need for it, as she would not want to "gate keep" the way students can use language to access and make sense of mathematics (e.g., the use of *slanted*). Instead, she would balance the push for language such that language functions as a resource with which she could utilise to create an environment where students can be comfortable and willing to use language to learn mathematics.

Additionally, based on how she attended to the meaning of the student discourse, rather than the specific words students used, Karen generally orientated towards *the use of language at the discourse level rather than word level*. For instance, when Ethel and Theo did not use the word *asymptote* to describe their observations, she was not overly concerned about their lack of the word. She was instead actively looking out for evidence showing how they were able to describe the characteristics of asymptotes in their own words. In particular, she commented, "it sounds to me like they know what's happening, they just don't have the word for it". It was, hence, a "clear-cut" situation for the term *asymptote* to show up. In her perspective, the learning and use of language has to be motivated by and *integrated within the learning of mathematics, rather than be an additive component*. As she would then involve the students in discussion to clarify their understanding of *asymptote,* in terms of the definition and the characteristics, it further suggests that Karen would likely place a strong focus on *developing conceptual understanding* in her mathematics classroom. In another task, which involved a subtraction of two integers, it was also noted that she did not immediately focus on correcting the procedures but considered how to clarify students' understanding of the procedures in a more conceptual manner.

A SMALL DISCUSSION

Through *accounting-for* how Karen would notice and use language through the lens of language-related dilemmas and orientations, it is evident that her primary concern with language in the mathematics classroom resides with students' use of language to acquire and develop understanding of mathematical concepts. Specifically, language functions as an important resource for mathematical thinking and learning, where both the mathematics register and students' everyday register play important roles. As such, she would strive to keep a good balance between the visibility and invisibility of language as a resource for mathematics teaching and learning. Using Adler's (2002) notion of a *transparent resource*, language is likely a *transparent resourc*e used by both Karen and her students to develop mathematical understanding in her classroom. Perhaps, as a final note for future research, it was interesting to observe how language tends to take a more secondary role as a resource, rather than a primary one in the mathematics classroom, to most (if not all) the teachers in my study. Notably, even Karen, who would deem language as an important resource for developing mathematical understanding, commented:

> it is my job to get them to be precise, but my first job is to get them to do math. And then my second job is to get them to be precise, so that they can communicate to an audience.

References

Adler, J. (2002). *Teaching mathematics in multilingual classrooms.* Springer.

Adler, J. (2021). Content and context specificity matter in the 'how' of language-responsive mathematics teacher professional development. In N. Planas, C. Morgan & M. Schütte (Eds). *Classroom research on mathematics and language: Seeing learners and teachers differently* (pp. 77–100). Routledge.

Halliday, M. (1975). Some aspects of sociolinguistics. In Final Report of the Symposium: *Interactions between Linguistics and Mathematical Education* (pp. 64–73). UNESCO.

Lampert, M. (1985). How do teachers manage to teach? *Harvard Education Review 55*(2), 178–194.

Mason, J. (2002). *Researching your own practice: The discipline of noticing*. Routledge.

Moschkovich, J. (2002). A situated and sociocultural perspective on bilingual mathematics learners. *Mathematical Thinking and Learning, 4*(2), 189–212.

Pimm, D. (1987). *Speaking mathematically.* Routledge & Kegan Paul.

Prediger, S., Şahin-Gür, D., & Zindel, C. (2019). What language demands count in subject-matter classrooms? *Research in Subject-Matter Teaching and Learning, 2*(1), 102–117.

Vygotsky, L. (1934/1986). *Thought and language* (A. Kozulin, Ed. & Trans.). MIT Press.

Zazkis, R. (2000). Using code-switching as a tool for learning mathematical language. *For the Learning of Mathematics, 20*(3), 38–43.

Zazkis, R., & Hazzan, O. (1998). Interviewing in mathematics education research: Choosing the questions. *The Journal of Mathematical Behavior*, *17*(4), 429–439.

A UNIDIMENSIONAL, MULTI-STRAND MEASURE VERIFIES A 6-SCHEME MODEL OF FRACTIONAL REASONING

Ron Tzur[1], Rui Ding[2], Bingqian Wei[1], Michael Sun[3], Beyza E. Dagli[1], Xixi Deng[2]

[1]University of Colorado Denver, [2]Northeast Normal University,[3]Englewood High School

We report on a new assessment to help address the problem: How may a feasibly large-scale written-test, informed by a 6-scheme constructivist model, measure students' fractional reasoning in different countries? We developed and validated the 35-item measure as a proxy of labor-intensive assessment forms of this model. We used mixed methods to develop it in English, translate it to Chinese, and analyze its properties to verify that (a) it is reliable (α >0.95) and valid (unidimensional) and (b) each scheme's items constitute a stand-alone, reliable (α >0.7) strand. We present initial findings of student responses (USA, n=61; China, n=217) that indicate similarities in their reasoning and discuss implications of our design, validation processes, and findings about students' fractional reasoning to theory, future research, and practice.

We address the problem: How may a large-scale written-test, informed by a 6-scheme constructivist model, measure students' fractional reasoning in different countries? Such an assessment tool is important in support of conceptually teaching and learning fractions – one of the most challenging areas in elementary mathematics (Lamon, 2007). Our measure builds on studies that stressed the hindering role that a part-of-whole concept of fractions may play in such learning in students (Tzur, 2019). Instead, we draw on an alternative of fostering reasoning about fractions as multiplicative relations, or measures (Simon et al., 2018). Specifically, we follow Tzur & Hunt's (2022) description of an 8-scheme progression of such reasoning (see Framework), itself a summary of a constructivist research program in this area (Steffe & Olive, 2010).

Commonly, inferring into students' reasoning involved qualitative methods, but this is labor intensive, with small samples. Recently, Wilkins et al. (2013) developed written forms to assess students' fractional reasoning. However, it requires an expert judgment of each response is indicative of the scheme being assessed. Training such experts and their work of scoring are, again, labor intensive. Thus, drawing on Kosko's (2019) item design for multiplicative reasoning, we examined the feasibility of measuring students' reasoning, at a large scale, based on their written responses as given.

THEORETICAL AND CONCEPTUAL FRAMEWORK

We draw on a constructivist stance (Piaget, 1985) and its core notion of assimilation into one's available schemes (von Glasersfeld, 1995). Our design of items followed Tzur and Hunt's (2022) description of an 8-scheme progression comprised of two, 4-scheme clusters, one based on iterating and the other on recursively partitioning fractional units. In our measure we focused only on the first 6 schemes (FR-6), and on

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. 4, pp. 169–176). PME.

a multiplicative reasoning (MR) with whole numbers. This choice reflects our findings with USA students who were yet to construct the two most advanced schemes. For compatibility, we kept this choice also in the Chinese version.

In the iteration-based cluster, first in the progression is an *equipartitioning (EP) scheme*, involving reasoning about unit fractions as a multiplicative relation (e.g., 1/3 is a unit that the whole is 3 times as much of it) and on the inverse relations among such units (e.g., 1/3>1/4 because 3<4). Second is a *partitive fraction (PF) scheme*, involving operating on composite fractions (e.g., 3/7+2/7=5/7, or 3/7*2=6/7). Third is an *iterative fraction (IF) scheme*, involving operating on any composite fraction (e.g., considering the effect of iterating 1/3 seven times as 7/3 of the "whole" and/or of iterating a unit of 3/3 twice + 1/3). Fourth is a *reversible fraction* (Rvrs) scheme, involving undoing the process that produced a composite fraction (e.g., partitioning 4/7 to produce 1/7 of the "whole," which is then repeated 7 times to reproduce the whole).

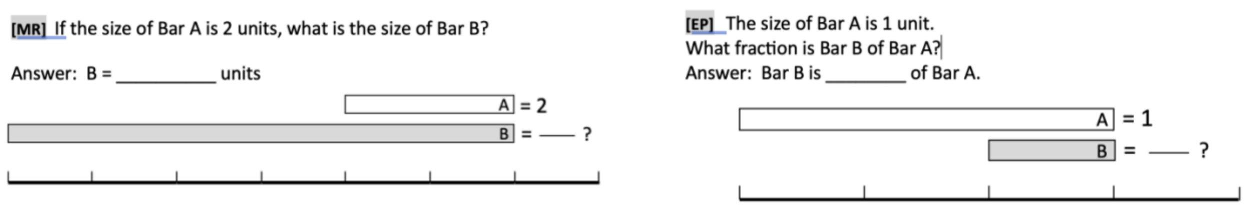
In the recursive partitioning cluster, fifth within the measure, is a *recursive partitioning* (RP) scheme, involving finding a unit fraction of a unit fraction of the "whole" (e.g., 1/3 of 1/5 is 1/15, because the whole is 15 times as much of it). Sixth is a *fraction composition* (FC) scheme, involving multiplicative coordination of unit and composite fractions (e.g., 1/3 of 2/5 of the whole is anticipated to produce two units, 1/15 each).

METHODS

This study was part of two larger projects (see acknowledgments). Here, we present the FR-6 measure (35 items) and report on the development and validation processes conducted with human-subject permission from the first author's institution.

A Measure of 6 Schemes in Fractional Reasoning (FR-6)

The first, full version of the FR-6 measure consisted of four items per fractional scheme, with four additional items to measure a related aspect of the EP scheme (inverse relations) and the RP scheme (equivalent fractions). A simpler version consisted of items to assess the first three schemes (FR-3). As our model stresses reasoning multiplicatively about fractional units, we also included three items to assess students' MR. All items were designed in line with Kosko's (2019) items for assessing MR. He used simple diagrams of relationships among units to minimize text to be read and avoid some of the issues associated with interpreting student responses to word problems. Figure 1 shows only one item for each scheme (two for Rvrs, initial and revised). Items for each scheme's subscale differ only in their numerical values. Based on our experts' feedback, we ordered items randomly, with one exception: to motivate students we chose the first three items to be the easiest (MR, EP, and PF).



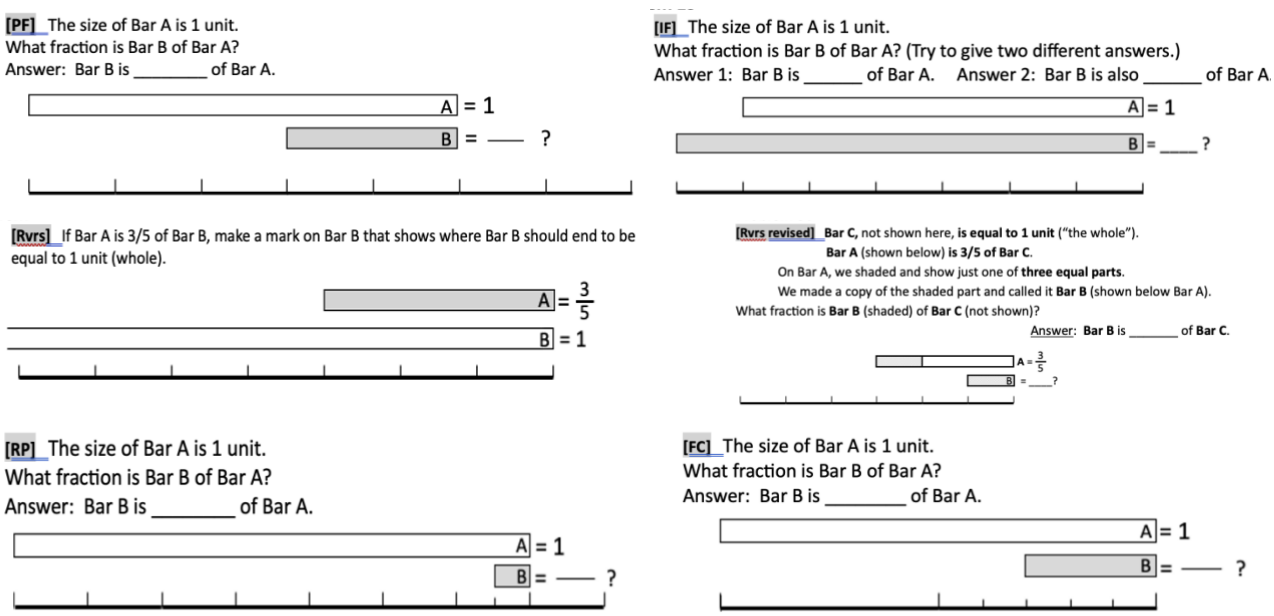


Figure 1: Examples of each scheme's items ([names] are not given in the measure).

Validity and Reliability

We used a 6-phase process to develop and validate the FR-6 measure in English. First, using tasks from prior research, the first author designed one item for each scheme. Second, based on 3 experts' feedback, he created all items for the MR and FR schemes, along with practice problems included in an "Assessor's Classroom Guide" (ACG). Third, he sent this full draft to an 8-expert panel, who responded with: (a) keep as is, (b) keep with reasoned changes, or (c) omit. The experts' feedback included no "omits" but many suggested changes to accentuate the way each item enables inferring the student's available scheme. Fourth, he used their feedback to create a full, first edition version. He used it (a) in a clinical interview with a participant of our teaching experiment and (b) to obtain further feedback from two experts in multi-language education. Fifth, he shared the ACG with teachers who administered the FR-6 measure. Sixth, he used Rasch modeling (Bond & Fox, 2015) to cement construct validity.

We then utilized a back-translation strategy to create a comparable FR-6 measure in Chinese. The second author (a Chinese native speaker) first translated the English version into Chinese. Next, the fourth author, *who never saw the original version* (a Chinese native speaker with BA and MA in the USA), translated the Chinese version back into English. The first author then compared the original and back-translated English versions, highlighting a few incompatible phrases. Last, the third author (a Chinese native speaker with PhD in the USA) checked highlighted phrases in both English versions to reconcile differences and finalize the Chinese version. This back-translation process guarantees the English and Chinese versions are equivalent.

Data Collection

We conducted our study with 4th-6th graders (n=217) in a high SES, urban, primary school in northeast China and 8th graders (n=61) in a low SES, urban, middle school in the southwest of the USA. Teaching fractions in both countries typically begins at grade 4. Yet, our work with those USA students indicated that the needed variance may be obtained from middle school students and, hence this difference in grade levels. Graduate research assistants (GRA), along with classroom teachers, used the ACG to administer the measure to all students present in a regular mathematics class.

Data Organization

To ensure reliable data organization, in each country we trained the GRA to work in pairs while entering the data (without identifiable information) into an Excel spreadsheet that automatically coded those entries. For IF responses we used “0” for incorrect, “1” for correct on mixed numbers only, “2” for correct on improper fractions only, and “3” for correct on both. For the other six schemes, we used “0” for incorrect and “1” for correct responses. First, we trained GRA while entering ~20% of student responses. When the GRA gained competence, they entered the remaining data. We then used those organized data as a basis for Rasch modeling, also creating an SPSS file that included both raw and Rasch person-ability scores to run all other analyses.

Data Analysis

We first used Winsteps 5.6.2 to conduct Rasch modeling of the FR-6, FR-3, and each scheme separately. This allows converting ordinal scales (raw scores) to interval scales that (a) warrant using parametric statistics and (b) assist in determining the properties of a measure (Bond & Fox, 2015). Using the Rasch modeling, we determined reliability using its Cronbach’s $\boldsymbol{\alpha}$ values and other properties (e.g., dimensionality). We also used (a) principal component analysis (PCA) to find if each scheme can be considered as a stand-alone strand (factor), (b) PROCESS macro for SPSS (Hayes, 2021) to run mediation (linear regression) analysis of the effect of early schemes on upper-level ones, and t-tests (SPSS) to find if subscale differences are statistically significant.

RESULTS

Reliability is a necessary but insufficient condition for construct validity (Nunnally & Bernstein, 1994). We thus begin with reliability analysis, then use Rasch modeling and PCA to establish unidimensionality of the measure and of each scheme. Finally, we provide analyses of student responses that seem common to both countries.

FR-6 Measure Consistency Analysis

Using Cronbach’s α from Rasch modeling, we found that in both countries the entire FR-6 measure (35 items), the FR-3 measure (16 items), and each scheme (4 items) separately have high alpha values (0.77-0.96), except for MR (see Table 1).

	FR-6	FR-3	MR	EP	PF	IF	Rvrs	RP	FC
China	.96	.91	.53	.81	.81	.81	.85	.90	.90
USA	.95	.88	.65	.77	.90	.91	.89	.93	.86

Table 1: Cronbach's α of the FR-6 and FR-3 measures and of each scheme.

Analysis of each scheme's Cronbach's α if an item is eliminated indicated this is possible for all schemes except for Rvrs. It also showed the need to add an item to the MR scale. We revised the measures accordingly and, so far, administered it to another group of 7th graders (n= 53) in the USA. Table 2 shows the revised α-values (# of items in parentheses; for the 13 items of the FR-3 subscale, α=.83).

	FR-6 (29)	MR (4)	EP (3)	PF (3)	IF (3)	Rvrs (4)	RP (3)	FC (3)
USA	.91	.71	.85	.86	.78	.69	.86	.76

Table 2: Cronbach's α of the FR-6 measure and each scheme (post-revision, USA).

Rasch Modeling

First, we found that variances explained by the measure in China and the USA, respectively, are 45.4% and 43.0% of the total variance – nearly the same as the model-expected variances of 45.1% and 42.9%. Each of the three and five contrasts in China and the USA, respectively, which show unexplained variances, was found to include scheme subscales. Second, all items show mean square (MNSQ) Infit and Outfit values between 0.5-1.5, which are considered productive for measurement, and all Point-Measure correlations are positive and large (USA: $\geq$ +0.57; China: $\geq$ +0.46).

Principal Component Analysis (PCA)

Using PCA in each country (SPSS), with KMO and Bartlett's tests (> .9, p < .001) and high communality levels (> .6), indicated the likelihood of finding such sub-dimensions. Unrotated component matrices and scree plots showed three components with eigenvalue > 1.0 in China and five in the USA. For each scheme separately, PCA indicated that it includes only one component with eigenvalue > 1.0, and all item loadings on each component were high (> .7). Along with the Rasch modeling, these findings led us to conclude that the FR-6 is a *unidimensional, multi-strand measure* – with each scheme being a stand-alone, unidimensional strand (subscale).

Similar Patterns in Student Responses

Using Rasch difficulty levels, we noticed a pattern common to Chinese and USA participating students (Figure 2). It corroborates scheme theoretical ordering within the iteration-based cluster – EP has the lowest difficulty-level (33.2, 35.9), then PF (45.4, 48.2), and then IF (69.2, 55.7), or within the recursive partitioning cluster - RP (47.9, 57.5) is lower than FC (59.1, 64.4). Independent samples t-tests between each pair of those schemes showed they are all statistically significant.

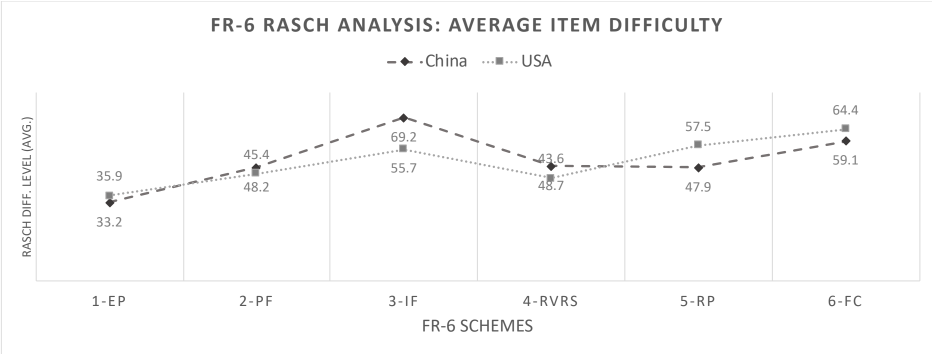


Figure 2: Rasch difficulty-level of each scheme.

However, it does not corroborate the conjecture that Rvrs is more difficult than IF, with the issue seemingly being Rvrs item design. Our initial (first edition) design of items for this scheme likely allowed students who had the EP or PF schemes to correctly respond to each Rvrs item. We thus redesigned the Rvrs items. Early findings from using the revised FR-6 (USA) indicated that it improved the construct validity of the Rvrs subscale, but it is still easier than IF (see Discussion).

Concept-Path Regression Analysis

To further corroborate the theoretical model, we conducted linear regression of the extent to which lower-level schemes predict more advanced ones. Figure 3 presents this concept-path "map," with standardized β coefficients entailing the slope in the corresponding linear equation (we show a path in China, as it lends further support to the Western-born model). For example, β-values of EP as a predictor of PF (0.7), IF (0.4), Rvrs (0.61), RP (0.72), and FC (0.56) are all statistically significant at $p<.001$.

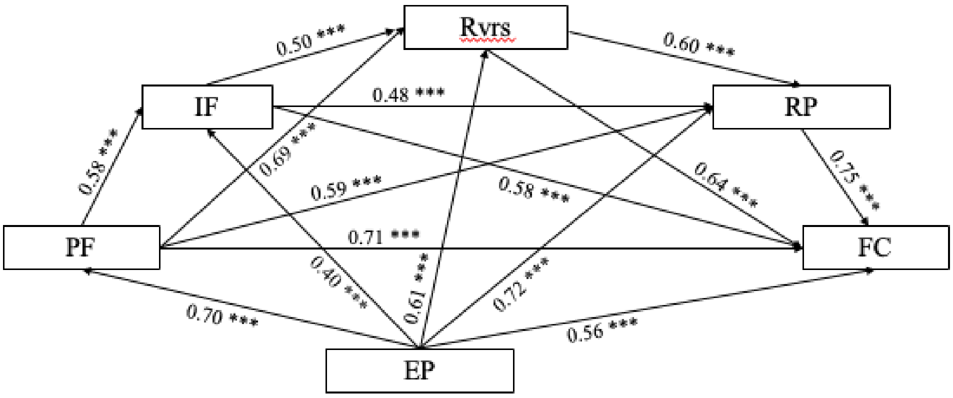


Figure 3: Concept-path of linear regression β-values (*<.05, **<.01, ***<.001).

DISCUSSION

Our study, in two different countries (China, USA), showed that this measure can serve as a feasible, valid and reliable, large-scale written measure of the first six schemes of fractional reasoning. It supported this claim for the entire FR-6 measure, for the first three schemes (FR-3), and for each scheme as a stand-alone strand. Assessors can thus use these eight scales as proxies for interviewing and/or using measures of open-ended items that require labor intensive scoring and training. Below, we discuss implications of our findings for theory building and future research, and for practice.

Implications for Theory Building and Future Research

Our findings (Figures 2 and 3) seem to corroborate the 6-scheme (out of 8 in two clusters) progression in fractional reasoning (Tzur & Hunt, 2022). In the measure, the first cluster includes the Equipartitioning (EP), partitive fraction (PF), iterative fraction (IF), and Reversible (Rvrs) schemes; the second includes the recursive partitioning (RP) and fraction composition (FC) schemes.

As for IF being more difficult than Rvrs, two conjectures can inform future studies. First, an issue could be the Rvrs item design. The ordering of IF and Rvrs has been identified through teaching experiments in which tasks for the Rvrs were much more challenging than tasks we used in the measures. In those experiments, a student would be given just a linear figure (no marks), told it is a fraction of the "whole" (e.g., 4/7), and asked to use it to produce the whole. In a written form, such a task entails the student has to produce a drawing of the "whole," which requires a person to score student responses – which contradicts a key goal of our measure development. Our revised tasks focused on recognizing the unit fraction that produced the given fraction (e.g., one of 4 parts is 1/7 of the whole). These tasks miss the second part of the scheme – producing the whole by iterating the 1/7. Moreover, they provide marks that allow a student to use EP or PF in responding to the item. Second, at least for some students with a strong PF scheme, solving Rvrs tasks may be conceptually more accessible than to reason with the IF.

Implications for Practice

The original, and the revised (shortened) versions of the FR-6 and FR-3 measures, have been administered and scored by teachers in our projects, following a classroom slide-show guide (the ACG) and a spreadsheet into which they entered student responses "as is." The spreadsheet automatically presents student outcomes at an individual and classroom levels. For practice, our study thus implies that teachers can use these measures, as well as each scheme's subscale, not only as a summative but also as a formative form of assessment. Specifically, once students are assessed on one of these measures, a teacher can (a) identify the scheme they seem to have, (b) determine the scheme to foster next, and (c) use only the subscale of the next scheme to assess each student's construction of it. Of course, until further research clarifies the above theoretical issue, assessing the Rvrs scheme may require interviewing students who could solve the tasks but are yet to construct this scheme.

Acknowledgement

This study was supported by grants from Sheridan School District #2 and from the Entrusted General Education Research Project of the Chinese Society of Education (2021). The opinions expressed do not necessarily reflect the views of the funders.

References

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences* (3rd ed.). Erlbaum.

Hayes, A. F. (2021). *Introduction to mediation, moderation, and conditional analysis: A regression-based approach* (3rd ed.). The Guilford Press.

Kosko, K. W. (2019). A multiplicative reasoning assessment for fourth and fifth grade students. *Studies in Educational Evaluation*, *60*, 32-42.

Lamon, S. J. (2007). Rational numbers and proportional reasoning: Toward a theoretical framework. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 629-667). Information Age.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory*. McGraw Hill.

Piaget, J. (1985). *The equilibration of cognitive structures: The central problem of intellectual development* (T. Brown & K. J. Thampy, Trans.). The University of Chicago. (1975 (French))

Simon, M. A., Placa, N., Avitzur, A., & Kara, M. (2018). Promoting a concept of fraction-as-measure: A study of the Learning Through Activity research program. *Journal of Mathematical Behavior*, *52*, 122-133.

Steffe, L. P., & Olive, J. (2010). *Children's fractional knowledge*. Springer.

Tzur, R. (2019). Developing fractions as multiplicative relations: A model of cognitive reorganization. In A. Norton & M. W. Alibali (Eds.), *Constructing number: Merging perspectives from psychology and mathematics education* (pp. 163-191). Springer Nature.

Tzur, R., & Hunt, J. H. (2022). Nurturing fractional reasoning. In Y. P. Xin, R. Tzur, & H. Thouless (Eds.), *Enabling mathematics learning of struggling students* (1st ed., pp. 315-335). Springer - Nature.

von Glasersfeld, E. (1995). *Radical constructivism: A way of knowing and learning*. Falmer. (1995)

Wilkins, J. L. M., Norton, A., & Boyce, S. J. (2013). Validating a written instrument for assessing students' fractions schemes and operations. *The Mathematics Educator*, *22*(2), 31-54.

GRUNDVORSTELLUNGEN IN UNIVERSITY MATHEMATICS – THE DEFINITION OF THE LIMIT OF A SEQUENCES

Karyna Umgelter and Sebastian Geisler

University of Potsdam

In this paper, we analyse the presentation of the definition of the limit of a sequence using the theory of Grundvorstellungen. Grundvorstellungen are mental images that lie behind mathematical concepts and support the development of valid concept images. The sample consist of six definitions presented by six different lecturers at German universities. The results show that lecturers usually address at least one Grundvorstellung when introducing the definition of the limit of a sequence. However, it is questionable, if this is enough to form a coherent concept image of the limit of a sequence. Finally, we give implications for further research.

INTRODUCTION

It has long been a known fact that many students struggle with studying mathematics and, as a result, often drop out from their mathematics study programs during their first year at university (Geisler, 2020). In particular, researchers criticise lecturing as a common way of teaching mathematics (e.g., Paoletti et al., 2018). Moreover, Viirman (2021) indicates the lack on research on teaching practices in mathematics lectures. For a better understanding of teaching actions of lecturers in mathematics lectures, Fukawa-Connelly (2014) emphasizes the need for new theoretical approaches. Especially the presentation of definitions as a central part of mathematics lectures should be investigated in more detail, also because several studies report that students struggle dealing with new concepts (e.g., Bills & Tall, 1998). In this paper, we are going to analyse the presentations of the definition of the limit of a sequence as it is a core topic in real analysis lectures with the focus on the theoretical approach of Grundvorstellungen. Grundvorstellungen are essential for robust understanding of mathematical concepts (e.g., Greefrath et al., 2021). In the following, we present the embedding of Grundvorstellungen in the theory of concept formation and summarize the results on presentation of definitions from previous research.

THEORETICAL BACKGROUND

Concept definition, concept image, mathematical aspects and Grundvorstellungen of mathematical concepts

Mathematical concepts form the core of mathematics (e.g., Halverscheid & Pustelnik, 2013). Tall & Vinner (1981) emphasize the importance of the distinction between formal definition of mathematical concepts and cognitive processes behind it. They describe *concept definition* as “a form of words used to specify [the] concept” (Tall & Vinner, 1981, p. 152). In the context of concept definition, Greefrath et al. (2021)

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. 4, pp. 177–184). PME.

describe mathematical definitions as a composition of different mathematical *aspects* of that concept. In other words, mathematical aspects emphasize different characteristics of concepts in various problem situations (Greefrath et al., 2016). They are the product of mathematical analysis of concepts (ibid.) and leave no room for interpretations (Greefrath et al., 2021).

Unlike concept definitions, *concept images* "describe the total cognitive structure that is associated with the concept, which includes all the mental pictures and associated properties and processes" (Tall & Vinner, 1981, p. 152). Tall and Vinner (1981) draw attention to the fact that students' concept image can conflict with the concept definition, therefore it is a major challenge to contribute to the development of correct concept images to get full understanding of a concept definition. In this sense, Fukawa-Connelly and Newton (2014) underline the importance of presentation of informal content in mathematics lectures like examples. Moreover, Capaldi (2020) points to the fact that lecturers' activities concerning presentation of definitions influence students' concept images. Therefore, students could develop different concept images related to the same definition of a concept in courses given by different lecturers (Capaldi, 2020).

Grundvorstellungen (or GVs, for short), also called *Basic Mental Models*, are intended to help with the development of viable concept images (Greefrath et al., 2021). They provide substantive interpretation of a concept (Greefrath et al., 2016) and "are prerequisites for dealing with mathematical concepts in an insightful way" (Greefrath et al., 2021, p. 650). Mathematical aspects can form the basis for a specific GV and several GVs can be assigned to a specific mathematical aspect (Greefrath et al., 2016). The key difference between concept images and GVs is that Greefrath et al. (2016) distinguish GVs in *normative* GVs that are the product of didactical analysis of mathematical concepts, and *individual* GVs that are the product of individual learning processes and can be a part of concept images (Greefrath et al., 2021). In the theory of Tall and Vinner (1981), this specification does not occur. For successful learning processes it is important to bring individual GVs closer to normative GVs, whereby several GVs are necessary for a broad understanding of a mathematical concept (Greefrath et al., 2016). Because GVs are central for development of robust concept images, the presentation of different visualizations and examples is necessary.

Presentation of definitions in mathematics lectures

Although the research on teaching actions of lecturers in mathematics lectures at universities based on observations of actual teaching is scarce (e.g., Viirman, 2021), there are several empirical studies on presentation of mathematical concepts. Nevertheless, we were unable to find research on presentation of definitions with a focus on GVs in mathematics lectures. In the following, we summarize the research on presentation of formal and informal representations of definitions that are substantial for the development of helpful individual GVs that are in line with normative GVs.

Overall, results of observational studies indicate large differences in the presentation of concepts by different lecturers. Based on observations of 11 lectures by 11 lecturers, Paoletti et al. (2018) noticed that only one lecturer did not present any formal definition in his lecture. Observing 11 advanced mathematical lecture courses, Fukawa-Connelly et al. (2017, p. 577) conclude that "[i]nstructors present informal content, including examples, informal representation, [...] during their advanced mathematics lectures, at least some of the time". Observing the lectures by one lecturer, Essien (2014) indicates that the lecturer introduced formal definitions without any motivation and presented only one or two examples to each of it. Although the studies report about the presentation of formal and informal representations of definitions, it is not clear which or if any GVs were addressed by lecturers.

Based on observations of mathematics lectures by three different lecturers, Chorlay (2022) presents deeper analysis of the presentation of the definition of the limit of a sequence. Overall, he found great differences in motivation of the concept, presentation of formal definition, examples, and illustrations: only one lecturer worked out the need of defining limits of sequences; two lecturers gave well thought out informal definition, used illustrations and demanding examples. All lecturers presented a formal definition introducing various parts of the definition in non-linear order and suggested a specific ε when working on examples although ε should be arbitrary. One lecturer gave an informal definition leading to the development of misconceptions that could be the reason for the development of individual GVs that are not in line with normative GVs.

Cottrill et al. (1996) report about difficulties in understanding and dealing with the formal definition of the limit of a sequence among students. Typical misconceptions regarding the limit of sequences are the confusion of the limit of the sequence with its value because of the lack of understanding of infinite processes (Cottrill et al., 1996), and the conviction that the infinite process itself is a limit (Vinner, 1991).

THE PRESENT STUDY

Research question

Empirical research regarding teaching of mathematical concepts at universities, especially "based on observations of actual lecturing" (Viirman, 2021, p. 467) is rare. For this paper, we are going to analyse the presentation of definitions in advanced mathematics lectures with the focus on normative GVs using the definition of the limit of sequences as an example. Although there are several studies on concept images and concept definitions, the literature search did not produce any results concerning teaching of definitions with a focus on GVs at universities at all. Therefore, we want to answer the following research question: *What normative GVs do lecturers address to present the definition of a sequence and which differences in addressing these normative GVs can be identified between the lecturers?*

Mathematical aspects and Grundvorstellungen of sequences

Greefrath et al. (2016) state two mathematical aspects of sequences: the dynamic aspect and the static aspect. Considering the dynamic aspect, two points matter: "recognizing the possibility of construction of a null sequence" (ibid, p. 103), and "recognizing the limit or limit value, whereby "almost all" sequence elements are located" close to the limit value (ibid, p. 103). For the static aspect, "a follower element is sought from which all further follower elements lie in a given environment of an object" (ibid, p. 104). The static aspect is central for understanding the ε- n_0-definition of a limit (ibid.).

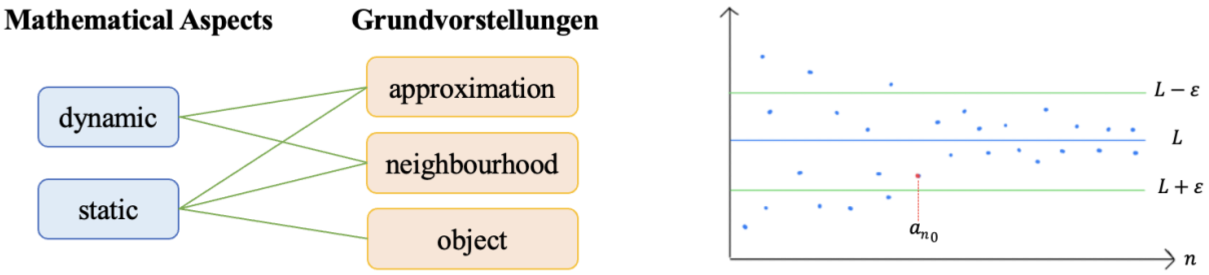


Figure 1: Aspects and Grundvorstellungen of the limit of a sequence (left) adapted from Greefrath et al. (2016, p. 97) and visual representation to the definition of the limit of a sequence (right), own representation.

Based on the mathematical aspects of the concept of limits (see Figure 1 left), Greefrath et al. (2016) identify three normative GVs: approximation, neighbourhood, and object. In the sense of the approximation GV, "striving towards or approximating of the sequence members' values to a fixed value or object provides the idea of approximation GV as an intuitive vision of the limit value" (ibid., p. 105). It should be mentioned that the sequence elements of an infinite convergent sequence $(a_n)_{n\in\mathbb{N}}$ come *arbitrarily close* to the limit L from a certain n_0 onwards. The neighbourhood GV (see Figure 1 right) is characterized by the following idea: "For every neighbourhood around the limit value, no matter how small it is, from a certain sequence element onwards all further elements lie in this area" (ibid., p. 105). It should be mentioned that the sequence elements of an infinite convergent sequence $(a_n)_{n\in\mathbb{N}}$ are in the ε-neighbourhood from a certain n_0 onwards. Moreover, n_0 depends on ε. Following the object GV, "limits are viewed as mathematical objects [...] that are constructed or defined by a sequence [...]" (ibid., p. 96). Thus, the limit of a sequence should be viewed as a mathematical object with which one can continue to operate.

Methodology

In this paper, we present the results of observations of the presentation of the definition of the limit of the sequence given by six different lecturers (we call them *Lecturer A*, *Lecturer B*, *Lecturer C*, *Lecturer D*, *Lecturer E*, *Lecturer F*) from five public universities in Germany: *Lecturer A* and *Lecturer B* gave their lectures at the same university for pure mathematics students and upper secondary pre-service teachers, *Lecturer D* and *Lecturer E* gave their lectures for upper secondary pre-service teachers, *Lecturer C* gave the lecture for secondary pre-service teachers, and *Lecturer F* gave the lecture for pure mathematics students. The lecture given by *Lecturer E* was audio recorded and board presentations were photographed, the lectures given by other lecturers were video recorded. Lectures on real analysis for pure mathematics students and upper secondary pre-service teachers are quite similar at German universities regarding their content. Observing the presentation of the definition of the limit of the sequence, we coded whether normative GVs regarding the limit of the sequence (see above) were addressed by lecturers, and the way the GVs were addressed. To ensure the reliability of the coding, we discussed the results with another mathematics educator researcher until we reached consent.

RESULTS

Overall, we identified two different approaches to the presentation of a formal definition of the limit of a sequence. For the first approach, the definition of the limit of a sequence given by *Lecturer A*, *Lecturer B*, *Lecturer C*, and *Lecturer D* can be summarized as follows: if for each $\varepsilon > 0$ there is an n_0 and $|a_n - L| < \varepsilon$ for $n \geq n_0$, the sequence $(a_n)_{n \in \mathbb{N}}$ is convergent and L is a limit of the sequence a_n. For the second approach, *Lecturer E* and *Lecturer F* used the definition of the limit of zero sequences: if $(a_n - L)$ is a zero sequence, a sequence $(a_n)_{n \in \mathbb{N}}$ is convergent. Next, we analyse the presentation of the definition by each lecturer with the focus on addressed GVs.

Lecturer A did not address the approximation GV while presenting the definition of the limit of a sequence. Instead, we observed the neighbourhood GV. *Lecturer A* presented a sketch (similar to the right part of figure 1) of an arbitrary convergent sequence $(a_n)_{n \in \mathbb{N}}$ a number line and explained how sequence members behave on the interval $(L - \varepsilon, L + \varepsilon)$ starting from an n_0. Moreover, he explained that n_0 depends on ε, the interval $(L - \varepsilon, L + \varepsilon)$ is arbitrary small and contains “almost all” sequence elements. In addition, *Lecturer A* characterised divergent sequences through an infinite number of sequence elements that are outside the interval $(L - \varepsilon, L + \varepsilon)$. Then, he presented several examples and nonexamples of convergent sequences: $(a)_{n\in\mathbb{N}}$, $\left(\frac{1}{n}\right)_{n\in\mathbb{N}}$ and ${(-1)^n}_{n\in\mathbb{N}}$. In case of the sequences $(a)_{n\in\mathbb{N}}$ and ${(-1)^n}_{n\in\mathbb{N}}$, *Lecturer A* used one sketch each to explain the behaviour of the sequence members on the interval $(L - \varepsilon, L + \varepsilon)$. We could not observe any GV addressed by the presentation on the example $\left(\frac{1}{n}\right)_{n\in\mathbb{N}}$. Concerning the object GV, *Lecturer A* presented only the notation of the limit of convergent sequences $\lim_{n \to \infty} a_n = L$ as a part of the definition, but he did not discuss possible operations with the limit of the sequence in more detail. Thus, the object GV was only strived.

Lecturer B addressed the approximation GV using a sketch to explain the distance of the sequence elements to the limit L. He mentioned that the distance between a_n and L can be “as small as we wish” in case “we choose n_0 big enough”. Moreover, *Lecturer B* used another sketch to address the neighbourhood GV. He highlighted that for a certain ε, all sequence elements from a certain n_0 lie in the neighbourhood of L, and this applies to all ε. Like *Lecturer A*, *Lecturer B* also presented only the notation of the limit of convergent sequences $\lim_{n \to \infty} a_n = L$ as a part of the definition related to the object GV without any discussion of operations with the limit of the sequence.

Lecturer C first addressed the approximation GV using the example $\left(\frac{1}{n}\right)_{n\in\mathbb{N}}$. He stressed, that with increasing n, the sequence elements get closer and closer to 0. Next, *Lecturer C* added a sketch of the example $\left(\frac{1}{n}\right)_{n\in\mathbb{N}}$ to explain the ε-n_0-Definition defining $n_0 = 4$ for $\varepsilon = 0.5$ addressing the neighbourhood GV. For this example, n_0 was chosen unfavourable because $n_0 = 3$ would have been the smallest n_0 fulfilling the definition. Furthermore, he states that the smaller the interval, the larger n_0 should be chosen, and all sequence elements from a certain n_0 are in the neighbourhood of L. Then, he underlined, that L does not have to be an element of the sequence. Presenting two more examples of convergent sequences $\left(\frac{n+1}{n}\right)_{n\in\mathbb{N}}$ and $\left(\frac{3n+1}{10+5n}\right)_{n\in\mathbb{N}}$, *Lecturer C* demonstrated how n_0 can be calculated for different ε and emphasized several times that all sequence elements from the certain n_0 are in the certain ε-neighbourhood. We did not observe any presentations related to the object GV.

Lecturer D addressed the neighbourhood GV using the example $\left(\frac{1}{n}\right)_{n\in\mathbb{N}}$. She used a sketch of the sequence and define its limit $L = 0$ without any explanation although at this point the approximation GV would be useful to work out the limit value. Then, she addressed the neighbourhood GV adding the ε-neighbourhood and stating, that for all ε, a "sufficiently large" n_0 can be chosen so that all sequence elements with $n \geq n_0$ are in the ε-neighbourhood. *Lecturer D* has indeed introduced the concept of divergence as the opposite of convergence, but she did not explain it in detail. Instead, she briefly mentioned that rules will subsequently be developed with which the limits can be easily calculated. Mentioning this, the object GV was briefly addressed but not explained in more detail. Then, *Lecturer D* presented several examples and nonexamples for convergent sequences: ${(-1)^n}_{n\in\mathbb{N}}$, $(n)_{n\in\mathbb{N}}$, $(|q|^n)_{n\in\mathbb{N}}$ for $|q| < 1$ and $|q| > 1$. Only presenting the example $(|q|^n)_{n\in\mathbb{N}}$ for $|q| < 1$ and $|q| > 1$, she mentioned that limits can be added up and that is a part of the object GV. We did not observe that any GVs were addressed in other examples and the approximation GV was not addressed at all.

To define the limit of a sequence, *Lecturer E* used the definition of a zero sequence. Defining a zero sequence, he addressed the approximation GV and explained that the inequality $|a_n| < \varepsilon$ must apply to all ε, so that $|a_n|$ is getting "smaller and smaller", "goes towards" 0, and n_0 depends on ε. Next, he presented the examples $\left(\frac{1}{n}\right)_{n\in\mathbb{N}}$ and $(|q|^n)_{n\in\mathbb{N}}$ for $|q| < 1$ without striving any GVs. Later, *Lecturer E* defined the limit of a sequence using the definition of a zero sequence, mentioned that a convergent sequence "strive towards a certain number", and presented the notation $\lim\limits_{n\to\infty} a_n = L$. Here, we identified the approximation GV, and the object GV as the use of zero sequences for the definition of convergent sequences. Finally, *Lecturer E* defines divergence as the opposite of convergence and give an example ${(-1)^n}_{n\in\mathbb{N}}$ without using any GVs. We could not observe the neighbourhood GV at all.

Lecturer F defined the limit of a sequence also using the definition of a zero sequence. Based on the neighbourhood GV, she used a sketch of the ε-neighbourhood and mentioned, that for a certain ε all elements from a certain n_0 are, “maybe not from the start”, in the neighbourhood of 0. Also using the example $\left(\frac{1}{n}\right)_{n\in\mathbb{N}}$ she stressed that the search is for an n_0 from which all sequence elements are in the ε-neighbourhood and demonstrated how n_0 can be calculated for different ε. Later, *Lecturer F* used the definition of a zero sequence to define the limit of a sequence, presented the notation $\lim_{n\to\infty} a_n = L$ and gave an example $\left(\frac{(-1)^n}{n^2}+5\right)_{n\in\mathbb{N}}$ stating that limit is 5; these correspond to the object GV as described above.

CONCLUSION AND DISCUSSION

Researchers point to the lack of empirical research on mathematics lectures based on observations (e.g., Viirman, 2021) and highlight the need for new theoretical approaches (Fukawa-Connelly, 2014). Therefore, we observed the presentation of the definition of the limit of the sequence given by six lecturers using the theory of Grundvorstellungen. The results show that the limit of the sequence was defined in these lectures through the ε-n_0-Definition-approach or through the definition of a zero sequence. In total, we observed the approximation GV by four lecturers and neighbourhood GV by five lecturers; three lecturers used both GVs. The object GV were observed by five lecturers, but it was not addressed in detail. Moreover, the example $\left(\frac{1}{n}\right)_{n\in\mathbb{N}}$ and the nonexample $(-1)^n{}_{n\in\mathbb{N}}$ were presented by several lecturers and seem to be common. As the lecturers did not address all normative GVs regarding the limit of the sequence in their lectures, their students could develop insufficient concept images concerning this concept. Students also could develop different concept images for the concept of the limit of a sequence due to the lecture they attend.

There are some limitations in or study. As our sample is not large, further observations are necessary. Furthermore, we observed only the definition of the limit of the sequence using the GVs, therefore research on presentations of further definitions and related normative GVs is needed. We expect that the object GV would be treated in more detail with the presentation of further content concerning sequences. Also, the development of students’ individual GVs should be investigated. Nevertheless, our research shows that it is possible to observe definitions presented in mathematics lectures regarding the addressed GVs. As normative GVs have not been worked out for all concepts relevant in university mathematics, more theoretical research is necessary. In addition, we recommend specific training for lecturers because most of them have no pedagogical training at all and are therefore not aware of the role of GVs for teaching and learning concepts.

References

Bills, L., & Tall, D.O. (1998). Operable definitions in advanced mathematics: The case of the least upper bound. *Proc. 22th Conf. of the Int. group for the Psychology of Mathematics Education* (Vol. 2, pp. 104-111). PME.

Capaldi, M. (2020). What definitions are your students learning? *Problems, Resources, and Issues in Mathematics Undergraduate Studies, 30*(4), 400-414.

Chorlay, R. (2022). Accounting for the variability of lecturing practices in situations of concept introduction. *International Journal of Mathematical Education in Science and Technology, 53*(5), 1071-1091.

Cottrill, L., Dubinsky, E., Nichols, D., Schwingendorf, K., Thomas, K., & Vidakovic, D. (1996). Understanding the limit concept: Beginning with a coordinated process schema. *Journal of Mathematical Behaviour*, *15*, 167-192.

Essien, A.A. (2014). Examining opportunities for the development of interacting identities withing pre-service teacher education mathematics classrooms. *Perspectives in Education, 32*(3), 62-77.

Fukawa-Connelly, T.P. (2014). Using Toulmin analysis to analyse an instructor´s proof attitudes toward mathematics. *International Journal of Mathematics Education in Science and Technology, 45*(1), 75-88.

Fukawa-Connelly, T.P., & Newton, C. (2014). Analyzing the teaching of advanced mathematics courses via the enacted example space. *Educational Studies in Mathematics, 87*, 323-349.

Fukawa-Connelly, T.P., Weber, K., & Mejia-Ramos, J.P. (2017). Informal content and student note taking in advanced mathematics classes. *Journal for Research in Mathematics Education*, *48*(5), 567-579.

Geisler, S. (2020). Early Dropout from University Mathematics: The Role of Students´ Attitudes towards Mathematics. In M. Inprasitha, N. Changsri, & N. Boonsena (Eds.), *Proc. 44th Conf. of the Int. group for the Psychology of Mathematics Education* (Interim Vol., pp. 189-198). PME.

Greefrath, G., Oldenburg, R., Siller, H.-S., Ulm, V., & Weigand, H.-G. (2016). *Didaktik der Analysis: Aspekte und Grundvorstellungen zentraler Begriffe*. Springer Spektrum.

Greefrath, G., Oldenburg, R., Siller, H.-S., Ulm, V., & Weigand, H.-G. (2021). Basic mental models of integrals: theoretical conception, development of a test instrument, and first results. *ZDM Mathematics Education*, *53*, 649-661.

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

Paoletti, T, Krupnik, V., Papadopoulos, D., Olsen, J., Fukawa-Connelly, T., & Weber, K. (2018). Teacher questioning and invitations to participate in advanced mathematics lectures. *Educational Studies in Mathematics, 98*(1), 1-17.

Tall, D., & Vinner, S. (1981). Concept image and concept definition in mathematics with particular reference to limits and continuity. *Educational Studies in Mathematics*, *12*(2), 151-169.

Viirman, O. (2021). University mathematics lecturing as modelling mathematical discourse. *International Journal of Research in Undergraduate Mathematics Education*, *7*, 466-489.

Vinner, S. (1991). Models of limit held by college calculus students. *Journal for Research in Mathematics Education*, *22*, 219-236.

STRESS MATTERS? A CORRELATIONAL AND EXPERIMENTAL STUDY ON THE IMPACT OF STRESS ON FRACTION NUMBER LINE ESTIMATION

Wim Van Dooren and Jordy Heusschen

Centre for Instructional Psychology and Technology, University of Leuven, Belgium

We investigated the impact of stress induction on the accuracy with which upper primary school learners conduct a fraction line number estimation task. The accuracy was investigated in a stress free and stress-induced condition, and reported stress levels were compared across conditions. A distinction was made between learners who are considered average mathematics performers as opposed to weak mathematics performers. Overall, stress induction led to lower accuracy, both for average and weak learners, while weak learners experienced a stronger increase of stress due to stress induction. Implications are discussed.

THEORETICAL AND EMPIRICAL BACKGROUND

Rational number sense through number line estimation

A well-developed number sense is crucial for the later learning of mathematics, and for functioning in today's society (e.g., Rittle-Johnson et al., 2016). One of the most challenging hurdles for learners regarding number sense is to develop a good understanding of rational numbers. They have been found to be substantially more difficult to understand than natural numbers, even for adults (Vamvakoussi & Vosniadou, 2004; Van Hoof et al., 2017). A common category of difficulties relates to the differences between rational and natural numbers. The natural-number prior knowledge interferes with the learning of and reasoning about rational numbers. Learners for instance think that 0.53 is larger than 0.7 because it is longer (a technique that would work for natural numbers), or that 8/13 is larger than 4/5 because 8 and 13 are larger than 4 and 5 (Vamvakoussi et al., 2018).

Researchers generally assume that numbers are represented on a mental number line, ordered left-right and from small to large (e.g., Booth & Siegler, 2006). In line with this idea, research on numerical magnitude understanding often uses number line estimation (NLE) tasks (Schneider et al., 2018), and most often the number-to-position task whereby a segment of the number line with a beginning and endpoint are given and a specific target number needs to be positioned. The accuracy of the estimates is then an indication of the numerical magnitude representation. Also in number line estimation, the estimation of rational numbers (and particularly fractions) has been found particularly difficult. Iulcano and Butterworth (2011) found that both children and adults made less accurate estimates of fractions than of natural numbers and decimals. The current paper therefore takes a closer look at the estimation of fractions on the number line.

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. 4, pp. 185–192). PME.

In recent years, researchers have come to understand that estimates on a number line do not occur directly, but are highly strategy based. Ashcraft & Moore (2012), for instance, saw that accuracy is higher for numbers close to the endpoints and to the midpoint, suggesting that learners use benchmarks. Specifically for number line estimations involving fractions, Van Dooren (2023) documented the variety in strategies that learners use, and how the adaptive use of strategies results in better performance.

The impact of stress on mathematical performance

It is well-known that learners experience varying degrees of stress while performing mathematical tasks. Some may have a degree of mathematics anxiety (e.g., Ramirez et al., 2018). Stress has been shown to negatively affect the available working memory, and through this mechanism affect performance on various cognitive tasks (e.g. Luehti et al., 2008), including mathematical tasks (e.g., DeCaro et al., 2010). This is particularly so for mathematically challenging tasks, such as those involving rational numbers. Working memory has also been convincingly been related (both correlationally, LeFevre et al., 2010, and experimentally, Askhenazi & Shapira, 2017) to performance on (natural) number line estimation tasks.

Most research regarding the impact of stress on mathematical and other cognitive tasks has been conducted in (young) adults. Research in school-age children is rare. An exception is the study of MacKinnon McQuarry et al. (2012) that showed that six-year olds with mathematical learning difficulties perform worse on both working memory tasks and mathematical tasks, and that the reported stress level correlated with performance, both in the group with and without mathematical learning difficulties, whereby children with learning difficulties reported higher levels of stress than those without.

RATIONALE FOR THE STUDY

The current study wants to take further steps in understanding the impact of stress on mathematical performance in school-age children. In our study, we do not only want to investigate the correlation between experienced stress and performance; we also want to look at the relation in an experimental way, in order to establish whether the relation can be seen as causal. Thus, we will not only look at self-reports of experienced stress, but also experimentally induce stress by means of time pressure. The study will thus work in two phases: a stress-free phase and a stress-induced phase.

We also want to get a clearer view on whether the mathematical ability of learners acts as a moderating variable, because previous research has indicated that children with lower mathematical abilities tend to experience more stress which may be one of the causes of their weaker performance. Regarding mathematical performance, we want to focus on the estimation of fractions on a number line, as we know from previous research that this is a challenging task.

Detailed research questions and overall hypotheses will be presented after we have clarified the method and the design of our study.

METHOD

Sample

A total of 273 students from the 4^{th}, 5^{th} and 6^{th} grade of primary schools in Flanders (Belgium) and the Netherlands participated in the study, with a mean age of 10.26 years (SD = 0.95), with approximately equally many boys and girls. They came from 7 different schools and belonged to 16 class groups. Informed consent was obtained from the parents of all participating children, and the study protocol received ethical approval prior to execution.

For each learner, we obtained a measure of their overall mathematical ability, by looking at the results of a standardized mathematical achievement test. Based on percentile scores, this test divides learners to groups (from A to E), whereby groups D and E represent the weakest performing 25% of the students. In our sample, 22.2% of learners could be considered to have a weak mathematical ability, and they will be compared to 77.8% learners with an average mathematical ability.

Stress Scale

Learners filled in the stress scale at three time points in the study. The scale was based on a study by Meulman (2016), and consisted of a line of 20cm long divided in 5 intervals, respectively referring to “no stress”, “light stress”, “quite some stress”, “a lot of stress” and “a great lot of stress”. Children could mark their stress level anywhere on the number line. The stress scale was administered three times throughout the experiment: At the very start, after the stress free phase, and after the stress-induced phase. The measurement at the very start was a baseline measurement, allowing to accurately measure the impact of the stress caused by solving the number line estimation task in two different conditions.

Number line estimation task

Learners solved a number-to-position number line estimation task in which each number line had 0 and 1 as endpoints, and a total number of 30 fractions needed to be estimated: 15 in a stress free condition, and 15 in a stress-induced condition. Both subsets were counter-balanced, so that some learners got the first half of the fractions in the stress free condition and the other half in the stress-induced condition, and vice versa.

Within each subset, 10 fractions were considered as experimental items, and data analysis was only conducted on these items. They had double digit denominators and were considerably more difficult to estimate (e.g., 4/15; 9/19). The remaining 5 fractions had single-digit numerators and denominators (e.g., 1/3), and were considered as rather easy buffer items, aiming at keeping students motivated for conducting the task. They were mixed randomly with the experimental items.

Experimental procedure and conditions

After filling in the initial stress scale, learners solved the first set of 15 number line estimations under a stress free condition, implying that they were given all the time they needed to complete the task. Afterwards, learners had to indicate the amount of stress that they experienced while conducting that task.

Next, they were given another set of 15 number line estimations, but in a stress-induced condition. Stress was induced by imposing a time pressure: The time to estimate a given fraction was limited to 5 seconds. This was realised by allowing learners to turn to the next number line page only when the teacher announced this by a sound signal, and by again having students turn their page at the end of the 5 seconds. The 5 second time limit was determined in a pilot study as the median time in which learners estimated the fractions without time pressure. It was considered as a measure that would induce stress, but not to shorten the time so much that an appropriate estimation would be impossible.

RESEARCH QUESTIONS AND HYPOTHESES

The following research questions were put forward, each time with hypotheses.

RQ1: What is the impact of stress induction and mathematical ability on number line estimation performance, and is the impact of stress induction similar in average and weak learners?

> We expect that stress induction will lead to a lower performance as compared to a stress free situation, and we expect this effect to be larger in students with a weaker mathematical ability.

RQ2: Is there a difference in the reported stress level between average and weak learners, and is this the case both in the stress free and stress-induced condition?

> We expect that weak learners will report higher stress levels than average learners, and that these difference will be more pronounced in the stress-induced condition than in the stress free condition.

RQ3: Is there a relation between the reported stress level and performance on the number line estimation task?

> We expect that there will be a negative correlation between the reported stress level and performance on the number line estimation task. We additionally expect a negative correlation between the increase in stress level (by comparing the stress reported after completing a task with the baseline measurement) and performance on the number line estimation task.

RESULTS

Research question 1 – Number line estimation accuracy

Given that the reported stress levels substantially differed across grade levels (4^{th} grade M = 51.02, SD = 37.30; 5^{th} grade M = 50.00, SD = 27.83; 6^{th} grade M = 28.75, SD = 51.32), grade level was taken as a covariate in all analyses where stress level is a predictor.

The accuracy of learners' estimates on the number line was quantified by the most commonly used measure in the number line estimation literature, i.e. the percentage of absolute error (PAE). The fraction 2/11 is situated at 36.36 mm on the 200 mm number line. If a learner marks the fraction at 40mm, the PAE score is 1.82%, which is calculated as follows:

$$\text{PAE} = \frac{|40 - 36.36|}{200}\, x\, 100 = 1.82\%$$

Table 1 reports the accuracies (PAE scores) of average and weak learners in both conditions. There was a main effect of condition on the performance on the number line estimation task, indicating a higher performance (thus lower PAE) in the stress free condition than in the stress-induced condition ($F(1, 270) = 36.36$, $p < 0.001$ $\eta^2 = 0.064$). Average students performed better than weaker students ($F(1.270) = 152.88$, $p < 0.001$ $\eta^2 = 0.22$). Contrary to our expectations, the absence of an interaction between both indicates that the effect of stress induction was similar in weak learners and in average learners ($F(1.270) = 0.45$, $p = 0.50$ $\eta^2 = 0.001$).

Group	Stress free condition	Stress-induced condition	Average
Average learners (n = 211)	4.78 (4.64)	6.91 (5.39)	5.84 (5.52)
Weak learners (n= 60)	13.18 (11.99)	16.73 (12.74)	14.96 (11.37)
Total group (n = 279)	6.64 (7.77)	9.08 (8.64)	7.86 (8.32)

Table 1: Mean PAE of average and weak learners in the stress free and stress-induced condition (SD between brackets)

Research question 2 – Reported stress levels

First of all, we conducted a manipulation check to confirm that the reported stress level would be higher in the stress-induced condition than in the stress free condition. This was indeed the case: It was on average 25.83 (SD = 21.95) in the stress-free condition and 51.32 (SD = 31.98) in the stress-induced condition.

In the baseline measurement, weak learners already reported a somewhat higher stress score than average learners ($M_{weak} = 24.80$, $M_{average} = 29.47$). However, when looking at the increase in stress in the stress free phase as compared to the baseline, there was no difference in that increase between both groups (M_{weak} = 9.98, SD = 28.83, $M_{average}$ = 3.97, SD = 20.06, $t(270) = -1.51$; $p = 0.14$). When looking at the increase in stress from the base line to the stress-induction phase, however, in line with the expectations a significant difference between both groups was observed (M_{weak} = 34.55, SD = 28.85, $M_{average}$ = 22.91, SD = 25.67, $t(270) = -3.01$; $p < 0.01$), indicating that the increase in reported stress was much greater in the weak learners than in the average learners.

Research question 3 – Reported stress levels in relation to accuracy

Various correlations between the stress levels that are reported by learners and their accuracy in the number line estimation task can be considered. First of all, we saw a positive and significant correlation between the stress level reported at the baseline measurement point and the accuracy in the stress free and stress-induced phase (respectively $r = 0.13$, $p = 0.03$; $r = 0.21$, $p < 0.001$). Also the stress reported in the stress free phase and the accuracy in the stress free phase correlated positively ($r = 0.17$, $p = 0.05$), and finally also the stress reported in the stress-induced phase and accuracy in the stress-induced phase correlated positively ($r = 0.27$, $p < 0.001$).

Also the increase in stress levels can be correlated with accuracy. The increase in stress level from the baseline measurement to the stress free phase did not correlate with accuracy in the stress free phase ($r = -0.06$, $p = 0.30$), but the increase in stress from the baseline measurement to the stress-induced phase did correlate with the accuracy in the stress-induced phase ($r = 0.16$, $p = 0.01$).

CONCLUSIONS AND DISCUSSION

We investigated the impact of stress on the accuracy with which upper primary school learners conduct a challenging mathematical task, i.e. estimating fractions on a number line. The impact of stress was investigated both correlationally (by relating the reported stress level to the accuracy) and experimentally (by letting students solve the same task under a stress free condition and a stress-induced condition). We thereby looked at whether the effects would be different for learners who are considered of average ability as opposed to those who are considered to be weaker (as determined by a standardized mathematics achievement test).

First of all, it was observed that inducing stress (by putting students under a reasonable time pressure) indeed led to a lower accuracy on a number line estimation task. As such, this is not surprising given what is already reported in the literature for other tasks, and it is also not surprising that for a task where a precise, fine-grained and stepwise strategical approach is important (Van Dooren et al., 2023). But more importantly, we found that inducing stress was equally detrimental for accuracy in average than in weak students, while we had expected this effect to be stronger in weak students. It is possible that the stress induction that we implemented indeed was already rather strong even for average students, in a challenging task such as fraction number line estimation.

Second, regarding the stress levels that were actually reported, we indeed saw that they were substantially higher in the stress-induced condition than in the stress free condition, thereby providing a manipulation check. Weak learners start at a slightly higher stress baseline than average learners, but particularly in the stress-induced condition, stress levels reported by the weak learners are far higher.

Finally, we also observed correlations between the reported stress levels and accuracies, and between the increase in stress level due to stress induction and accuracies. Hence, where some learners experience a smaller stress increase due to the stress induction, they on average also suffer less from a decrease in accuracy than learners who experience a larger stress increase.

While this study is experimental and psychological in nature, and took place in a very controlled setting (although within the classroom), it does focus on a task that has curricular relevance (fraction number line estimation) and that is shown to be difficult for learners. The study has various limitations, such as the experimental setting which may be quite artificial, but this is also somehow a strength being able to show causal relations. A correct time pressure may have been different for different students (and different grade levels), but this was impossible to implement. Also the way in which stress was measured is open to criticism. Other ways of measuring this may be more qualitative (such as interviewing participants on their feelings during a task), but also physiological measures (such as skin conductance, pupil dilation) may be informative.

Understanding the impact of stress when conducting mathematical tasks is of great importance for practice, and the insight that weaker students may experience more stress than stronger students is crucial. Even while we were not able to show that the weaker students experienced a larger decrease in accuracy due to stress than average students, a stronger increase in stress as such is also undesirable in its own right. Time pressure may act differently for stronger and weaker students, and it may be that there are still large interindividual differences in sensitivity to time pressure, that do not boil down to the difference between weaker and stronger learners, but to other learner characteristics. This certainly deserves further research. Similarly, future research could investigate whether similar effects occur in other mathematical tasks.

References

Ashcraft, M. H., & Moore, A. M. (2012). Cognitive processes of numerical estimation in children. *Journal of Experimental Child Psychology, 111*(2), 246-267.

Ashkenazi, S., & Shapira, S. (2017). Number line estimation under working memory load: Dissociations between working memory subsystems. *Trends in Neuroscience and Education, 8–9*, 1–9.

Booth, J. L., & Siegler, R.S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology, 41*(1), 189-201.

DeCaro, M. S., Rotar, K. E., Kendra, M. S., & Beilock, S. L. (2010). Diagnosing and alleviating the impact of performance pressure on mathematical problem solving. *Quarterly Journal of Experimental Psychology, 63*(8), 1619–1630.

Iuculano, T., & Butterworth, B. (2011). Understanding the real value of fractions and decimals. *The Quarterly Journal of Experimental Psychology, 64*(11), 2088-2098.

LeFevre, J.-A., Fast, L., Skwarchuk, S.-L., Smith-Chant, B. L., Bisanz, J., Kamawar, D., & Penner-Wilger, M. (2010). Pathways to Mathematics: Longitudinal Predictors of Performance. *Child Development, 81*(6), 1753–1767.

Luethi, M. (2008). Stress effects on working memory, explicit memory, and implicit memory for neutral and emotional stimuli in healthy men. *Frontiers in Behavioral Neuroscience, 2,* 5.

MacKinnon McQuarrie, M. A., Siegel, L. S., Perry, N. E., & Weinberg, J. (2012). Reactivity to Stress and the Cognitive Components of Math Disability in Grade 1 Children. *Journal of Learning Disabilities, 47*(4), 349–365.

Meulman, A. F. (2016). *Het verband tussen gerapporteerde stress en gemeten fysiologische uitingen van stress tijdens rekentaken, verklaard door leeftijd en rekenprestatie*. (Masters thesis). Retrieved at http://dspace.library.uu.nl/handle/1874/327412

Ramirez, G., Shaw, S. T. and Maloney, E. A. (2018). Math anxiety: Past research, promising interventions, and a new interpretation framework. *Educational Psychologist, 53*(3), 145-164.

Rittle-Johnson, B., Fyfe, E. R., Hofer, K. G., & Farran, D. C. (2016). Early Math Trajectories: Low-Income Children’s Mathematics Knowledge from Ages 4 to 11. *Child Development, 88*(5), 1727–1742.

Schneider, M., Merz, S., Stricker, J., De Smedt, B., Torbeyns, J., Verschaffel, L., & Luwel, K. (2018). Associations of number line estimation with mathematical competence: A meta-analysis. *Society for Research in Child Development, 89*(5), 1467-1484.

Vamvakoussi, X., & Vosniadou, S. (2004). Understanding the structure of the set of rational numbers: A conceptual change approach. *Learning and Instruction, 14*(5), 453-467.

Vamvakoussi, X., Christou, K. P., & Vosniadou, S. (2018). Bridging psychological and educational research on rational number knowledge. *Journal of Numerical Cognition, 4*(1), 84-106.

Van Dooren, W. (2023). Strategy use in number line estimations of fractions - an exploratory study in search for adaptive expertise. *Proceedings of the 46th Conference of the International Group for the Psychology of Mathematics Education,* vol. 3, pp. 299-306

Van Hoof, J., Vamvakoussi, X., Van Dooren, W., & Verschaffel, L. (2017). The transition from natural to rational number knowledge. In *Acquisition of complex arithmetic skills and higher-order mathematics concepts* (pp. 101-123). London, England: Elsevier.

HUMAN GRAPHS AS MATHEMATICAL DRAMATIC CODIFICATIONS

Katherina von Bülow

Simon Fraser University

In this paper, critical educator Paulo Freire's theory and method of codification/decodification is applied, by means of a drama technique, to mathematics education. A classroom activity, in which students' bodies form a frozen tableau representing data on wealth disparity, is described. The study focuses on students' perceptions of their own relationship with the social issue and with mathematical representations of the data. Students' discussion and written reflections on the activity are analysed thematically. A parallel is drawn between shifts in students' recognition of themselves within the issue of social concern and shifts in their critical perceptions of mathematical representations of the data.

INTRODUCTION

Mathematics can serve to model and represent real data and patterns related to social justice issues. In mathematics education, we might create, re-create, interpret, change, and consume such mathematical representations. Critical researchers and educators ask how these mathematics activities work to position students with respect to the real social issue, mathematics, and themselves. In this paper, I consider the mathematical representation, through graphing, of numerical data that entails a social disparity concerning a population that includes the students themselves. Students may have different perceptions—or levels of awareness—of the social disparity, of their own relative part in this disparity, of mathematics as a possible way to connect with the issue, and of themselves as subjects in the mathematics classroom. My research aims to investigate and work with these levels of perception at the individual and collective levels. In this context, it is useful to think with Paulo Freire, who theorized education's work in the development of people's perception of reality and of themselves as subjects able to reflect and act on reality.

THEORETICAL FRAMEWORK

In *Pedagogy of the Oppressed* (1970/2000), Freire develops a theory and method of liberating, challenge-posing education. In Freire's method, educators first select elements of reality from students' "background awareness" (p. 83), where they exist objectively but are not yet objects of cognition and action. These must be elements that are present in situations that both involve basic contradictions and are part of students' lived experience. Each such situation is then "codified", that is, *abstracted* into an object, image, or short oral prompt that "shows some of [the situation's] constituent elements in interaction" (p. 105), in such a way that students will be able to recognize the situation, and to recognize themselves as subjects, together with other subjects, 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. 4, pp. 193–200). PME.

the codification (p. 114). The codification is presented to students, who collectively face it and are *challenged by it* to respond, so that "rather than students receiving information about this or that fact, they analyse aspects of their own existential experience represented in the codification" (Freire, 1985, p. 52). For example, Freire (1970/2000) recounts how a photograph of a local worker who is drunk functioned as a codification for a group to unveil the complexities of alcohol use and of a worker's experience within their community and context. Freire argues that, since students apprehend the challenge not as a theoretical question but within a context where the contradiction is interrelated with other issues, their responses tend to be critical and thus less alienated, and that in their responses they express their personal themes and view of the world. This reflexive response process, called *decodification*, stimulates "perception of the previous perception", "knowledge of the previous knowledge" (p. 115), and new perceptions and knowledge. Freire emphasizes that codifications should be "simple in their complexity and offer various decoding possibilities" (p. 115). It is in this unveiling process of decodification that "*generative themes*" (p. 96), that is, unfolding personal themes related to "key contradictions in people's lives" (Gutstein, 2012, p. 26), emerge, forming the curriculum for further investigation.

I am interested in the potential of drama for Freirean codification in mathematics education. As explained by Villanueva & O'Sullivan (2020), role-playing offers opportunities to reflect upon meanings that are created by one's own experience as well as collectively by the group. Thus, role-playing functions "as both an individual and socially reciprocal concept[, which] deepens its capacity for critical reflection and aligns it further with Freirean codification" (p. 529). In a dramatic codification, students are presented with a challenge related to a social situation, which as described above must involve a contradiction or disparity that is related to students' lived experience. The students act the situation out using a drama technique, such as role-play or frozen tableau, that is adapted to a classroom environment. Since one of the goals in this research is to investigate and work with perceptions of mathematics as a helpful partner in one's connection with social issues of concern, the situation under consideration must involve or have a link to mathematics. I submit that the link(s) between the action and mathematics can be explicitly clear from the beginning or be strong enough to emerge naturally in discussion afterwards. Henceforth, I will use the name MSJ (as in mathematics and social justice) dramatic codification to refer to this pedagogical method. Since students share the experience of being acting subjects both in the classroom and in the real-life situation, teachers and researchers can then invite students to reflect and express, through both discussion and individual writing, how they think or feel themselves in relation to the mathematics, the social context of the classroom, the social issue(s) of concern, and how these aspects are linked. As pointed out by Sean Chorney (personal communication, August 21, 2023), MSJ dramatic codifications can be both a teaching and a research method.

Circling back to the issue of mathematics representations, I am particularly drawn to dramatic codifications because they can centre elements, such as students' bodies, movement, emotional expression, and interaction, that can otherwise be under-represented in mathematics classrooms. In the pilot study I present here, students are asked to represent data showing the distribution of wealth in households in Canada, but to do so using their bodies, as a human graph, instead of using pencil and paper. There are similarities between this activity and Polly Kellogg's (2006) "Ten Chairs of Inequality" activity. However, rather than using chairs as a proxy for wealth, I am interested in maintaining the idea of graphing as a mathematical representation of data, while offering students the opportunity to question how graphs can be made. An unhelpful distance may be created between mathematics and the world when people, rather than representing data about themselves using their own bodies, use points or bars instead: the item in question is "not situated as a link in a chain of interaction between persons (to be understood and judged in communicative, social, and moral contexts)" (Barth, 1995, p. 65). Thus, the intention behind the use of physical mathematical representations—in this case a frozen tableau—is to counter a distancing tendency in how our social data might be presented or used in the classroom, and instead use a representation that aids students' awareness of their own involvement in the data. Another aim is to encourage a shift in students' perspective of mathematics as something with which we feel the world, not just think it, and as socially reciprocal, in that one can see others feeling it and know that they can see one feeling it.

Challenging students to use their bodies and the space between them in a mathematical representation parallels the challenge to their awareness of their own role in the social situation. Relative to the mathematics as to the social issue, students may shift, from submerged, alienated, adapted observers, to more aware, reflexive actors. This means that, from perhaps perceiving the data as simply statistics about finance, students could shift towards perceiving it as a social issue of concern to and about themselves and their peers; in parallel, perceptions of mathematics and of themselves within the mathematics classroom could shift towards including their own expression as well as investigation of issues that matter to them. Hence, my research questions ask:

- How does the human graph activity contribute to students' reflection about the social justice issue of wealth inequality?
- Do students reflect on mathematics and on themselves as mathematics students, and if so in what way are those reflections connected to the activity?
- Does the activity contribute to the emergence of students' generative themes?

METHODOLOGY

Context and Description of the Activity

This pilot study was conducted in a class from Dr. Chorney's Simon Fraser University course called 'Shape and Space', for liberal arts students. In class, the activity took approximately 45 minutes, and a questionnaire was given to students afterwards.

The activity began with a brief introduction that raised the issue of how things are distributed among people in Canada, and in particular, of how wealth is distributed. Students gave examples of different kinds of assets that could be included in the calculation of a person's total wealth. I then recounted how I searched Statistics Canada for information on wealth distribution and found wealth data for population quintiles, in terms of households, that is, persons or groups who share a dwelling, including temporarily absent members. The table below was displayed, at first with the actual wealth numbers missing so that students could predict the values A-E.

Canadian households ordered by wealth	Average wealth (net worth)
0 - 20%	A = - 2,762
20 - 40%	B = 125,936
40 - 60%	C = 429,271
60 - 80%	D = 946,048
80 - 100%	E = 3,139,492

Table 1: Distributions of Canadian household economic wealth

The average Canadian household wealth (measured in 2022) was revealed to be $927,597.00 (Statistics Canada). Before the actual values of A-E were revealed, students discussed where they would predict that average might fall, with respect to the letters A-E. After predictions were discussed, the numeric values A-E were given, and reactions and interpretations were shared. For instance, students discussed what a negative 'net worth' means. Students were then challenged to each take on a role (i.e., a wealth category) and form groups so that they could physically graph the data. Groups set out to find room, inside or outside the classroom, to represent this situation using their own bodies in space, rather than using points or bars on paper or board. Students simulated the situation by figuring out the relative distances at which to stand. The economic distances between people were symbolized and witnessed as spatial distance, and conversations ensued about how it feels, or must feel, to be in a specific position in that space of 'net worths'. Once all the students were back in the classroom, one of the groups was invited to show their process and their human graph. The group spread out diagonally across the classroom, and the whole class was challenged to make sure that the spatial distances between group members really corresponded to the numeric gaps in the data. Rounding and estimating were employed, until the class was satisfied that the bodies were graphing the data. Students remarked on how far from the rest the person playing the role of richest 20% Canadian households stood, while the remaining four group members stood comparatively close to each other.

Data Collection and Analysis

Here, I analyze the reflective writing of eighteen students (referred to as S1-S18) about the human graph activity, described above, responding to the following prompts:

- What were your impressions, as your group worked to decide how far apart to stand in order to represent the differences in wealth?

- How would you compare graphing this data on the board/paper to making a human graph? What information (if any) is lost/gained in either case? What kind of data do you think is suitable for either case?
- In this activity as a whole, how did you feel, physically and emotionally? What (if anything) felt helpful, interesting, unpleasant, or …?

In the analysis, students' responses are grouped into four themes corresponding to possible shifts in perception as theorized by Freire in the codification/decodification pedagogical method, adapted to mathematics education. For each theme, I give excerpts of student writing to illustrate the findings.

Theme: Connections between mathematics, oneself, and the issue(s) of social concern

Students share their awareness of being in a wealth category and some related feelings:

S18: I also started to think about where my family and I would land on this graph.

S2: …a humbling reminder of where I am, and likely will remain for the rest of my life. It's also nice to know that many, many others are there with me.

S13 I was slightly frustrated on a personal level considering I most likely fall within the lower/middle half of group "B", and how much accumulated wealth/assets people have in comparison to others.

As well, students share how the activity helped their understanding of the social issue:

S5: It was more surprising to see how A had the wealth of -3k. when thinking of the numbers aligning with the placements, I had not considered that a group would be in debt as I was thinking of positive numbers, not negative.

S18: This activity… enable[s] us to think about what determines someone's wealth and what being wealthy really means to everyone.

Theme: Feelings, beliefs, and physical sensations connected with the activity

Several students share that they felt "surprised", "shocked", "unpleasant", "disbelief", "envy", "saddened", "disturbed", "heartbroken", "disappointed", "jarred", "frustrated", and "sorrowful". For these students, the activity stirs feelings towards wealth disparity as a social issue of concern, and many share their awareness of its inherent contradictions. On the other hand, five of the students state they felt "fine", "neutral", "no feelings", or "normal". Here are some related excerpts:

S2: It brings out feelings of disbelief and envy. I feel an almost imperceptible "rumble" in my stomach area– one that wants to say, "is that even possible?!" and "is it possible to get there?!"

S7: Physically I felt quite far apart from my classmates, emotionally I felt quite sad to see the poorest category of 0-20% stand so far away from me.

S9: I felt very surprised throughout the activity because finding out how drastically different the two extremes were (0-20% and 80-100%) is heartbreaking to me.

S3: It doesn't seem right as just a small amount of the richest group could be spread out to the rest and would even out the poorer groups while the richest group would still be up there.

A longer study would report on further class discussion and investigation, for example asking what students mean when they state they feel "neutral" or when they ask, "is it possible to get there?!", continuing the dialogue begun by means of the codification.

Theme: Changes in feelings or beliefs about mathematics and about oneself as a mathematics student

Only two students make self-aware comments on "doing math" in the context of the activity:

S2: I definitely have a tendency to zone out and mindlessly do math. In an unfamiliar activity like this, that isn't likely to happen.

However, some students do take active critical positions with respect to mathematics and the activity. One student questions Statistics Canada's choice of categories:

S7: We are grouping large amounts of people into a 0-20% or 20-40% slot when in reality, the 18% from the first group and 22% from the second group could share very similar livelihoods, revenue, or net worth yet this won't be obvious because they will be separated into different categories.

Also, the reflections show students' engagement with the tension between privileging accuracy or emotional impact, when interacting with this data. I interpret this tension as concerning both the students' evaluations of the activity and their conceptions of mathematics:

S16: It is more time efficient to simply put up a slide of a graph to get your point across, but it is more memorable as a human graph.

S10: Making a human graph…can be a more visceral and impactful way to represent data, but it may sacrifice some of the precision and detail that can be conveyed through traditional graphs.

S13: Emotionally something about being physically present within the content material is valuable to one's connection to the material.

Notably, some students opine that a human graph is more "actual" or even "correct":

S3: Graphing on the board/paper may show you the use of real data and calculations in numbers which gives you an idea of what the data is but the human graph can let you visualize the actual differences.

S5: It was easier to move each person to their correct placement compared to the paper graph. Human graph helped us get closer results to the actual wealth numbers.

It is also noteworthy that students use several words that include not just intellect but also emotions and the body, to compare the human graphing with regular graphing. According to students, human graphing is or has:

- *more*: tangible, impactful, engagement, tactile feel, realistic, easy to move, visual, easy to grasp, perspective, effective, powerful, emotional impact, immediacy, visceral, drastic, emotional connection, eye opening, involved, thinking required, memorable; and
- *less*: objective, precision, overall adaptability, simple to do, accurate, exact, complex, detail, useful.

It would be interesting to have a follow-up class discussion about the meaning and relative importance of qualities they named, such as precision, efficiency, tangibility, realism, or immediacy, in relation to different choices of graphing for this data.

Theme: Interest in further investigation (possible generative themes)

Students' generative themes are present in their reflections, as in:

S6: I think that these huge disparities could be easily rectified through respective taxing based on income and pre-existing personal funds. I think that these disturbing numbers really showed me how much the government protects the upper class and disregards the impoverished.

S7: This put into perspective the unpleasant and frustrated feelings that this generation must feel towards the housing crisis and despite their hard work in their jobs, they still cannot afford what their parents could at their age.

S13: I think it would be more interesting to break down the data further, not just people across Canada but being more specific with one's socio-economic and geopolitical status, SES based upon one's race/place of living/gender… One step further would be to see and interview people from the "A" and "E" bracket and compare, I think that would cause one to be emotionally connected to their data.

S17: What kind of people will be lost (poorest people who are spending money on drugs, alcohols/poorest people who are spending money on their daily necessity)?

Discussion

The data from these eighteen students suggests that this human graph activity functioned as a dramatic Freirean codification/decodification. Firstly, the data may indicate a shift in students' ability to see themselves in the social issue of wealth disparity and to perceive it as a contradiction in their lives. Secondly, students' use of language involving feelings and physical sensations in describing their experience in the activity may indicate shifts in traditional conceptions of mathematics as neutral, context-free and related primarily to thinking. For example, students ponder the relative importance of effectiveness and precision, which are traditionally highly linked to mathematics, and other potentially desirable qualities, such as "visceral" or "memorable". As one student put it, in comparing regular graphing to the activity:

S2: It's like the difference between reading about the height of the Eiffel Tower, and standing at the base of it looking up.

This points to the start of a critical attitude from students with respect to classroom mathematical representations of their own data. Thirdly, because the activity challenged students to predict and then to compare their predictions to data from the 2022 Canadian census, it also helped students gain awareness of their previous perceptions and knowledge about the issue of economic disparity. For example,

> S6: I was surprised when we found out the actual depiction of these numbers because the actual spread was way larger (between 60-80% and 80-100%) than we had originally predicted.

Finally, the data shows students' generative themes, in which students propose a more detailed investigation of the data, of the experience of the people behind the data, of intersectionality with other social issues, and of policies such as taxation.

CONCLUSION

In the human graph activity described here, students took on roles from data categories on wealth, and walked away from each other to spatially represent those economic gaps. During the activity, mathematical objects were students' bodies, statistics were people's experience, and social concerns were emotions and voices in the classroom. The class discussion and written reflections show that students recognized themselves as part of this situation and of the disparity therein. Thus, students were aware of the mathematics they were doing as relevant to their own existential experience. Furthermore, students shared their views of the world and personal generative themes. I submit that the activity functioned as a mathematical dramatic Freirean codification/decodification, offering opportunities for shifts in students' perceptions of economic inequality, of their own part and position in this disparity, of mathematics as a possible way to connect with the issue, and of themselves as subjects who can critically analyse mathematical representations and mathematics classroom activities.

References

Barth, F. (1995). Other knowledge and other ways of knowing. *Journal of Anthropological Research, 51*(1), 65–58.

Freire, P (1985). *The politics of education: Culture, power, and liberation*. Bergin & Garvey.

Freire, P. (2000). *Pedagogy of the oppressed*. Continuum. (Original work published 1970).

Gutstein, E. (2012). Mathematics as a weapon in the struggle. In O. Skovsmose & B. Greer (Eds.), *Opening the cage: Critique and politics of mathematics education* (pp. 23–48). Sense.

Kellogg, P. (2006). Ten chairs of inequality. In E. Gutstein & B. Peterson (Eds.), *Rethinking mathematics: Teaching social justice by the numbers* (pp. 135–137). Rethinking Schools.

Statistics Canada (2023). Table 36-10-0660-01 Distributions of household economic accounts, wealth, by characteristic, Canada, quarterly (x 1,000,000) [Data table].

Villanueva, C., & O'Sullivan, C. (2020). Dramatic codifications: Possibilities and roadblocks for promoting critical reflection through drama in Chile. *Research in Drama Education, 25*(4), 526–542.

EARLY DIVISION PRIOR TO FORMAL INSTRUCTION: YOUNG CHILDREN EXPLAIN THEIR SOLUTION STRATEGIES

Luca Wiggelinghoff and Andrea Peter-Koop

Bielefeld University

This paper is embedded in a larger international study of young children's understanding of division prior to formal instruction. Real-world related division problems typically can be interpreted as either partitive or quotitive division and respective solution strategies. However, previous papers have identified children using grouping strategies that are typically related to a quotitive context for solving partitive division problems. The related drawings and written result do not explain the underlying thinking process. Hence, this paper focusses on the results of a qualitative study in which children were asked to explain their solution with or without modelling.

INTRODUCTION

Mathematics teaching as well as underlying curricula are frequently based on the assumption that young children develop a (formal) understanding of division predominantly in school. Furthermore, in many countries division is only introduced after the children have developed an understanding for addition, subtraction, and multiplication. At the same time earlier findings from studies carried out in the 1990s (Mulligan, 1992) as well as more recent research projects (Tumusiime et al., 2019; Cheeseman et al., 2022; Wiggelinghoff, 2022) suggest that many first and second graders can solve real world division problems successfully prior to the formal introduction of division in their mathematics classrooms. However, little is yet known about young children's conceptual understanding of division and its development. In this context, the two studies published by Tumusiime et al. (2019) and Wiggelinghoff (2022) seek to identify children's respective solution strategies and approaches. In their studies Grade 1 and Grade 2 students were asked to solve division problems presented through pictures in a paper and pencil test, that involved three partitive division and three quotitive division problems (see Fig. 1). The analyses of the drawings the children provided to show their answers, suggest that many children chose to solve partitive division problems by using a *grouping strategy* and hence use a strategy that works well for solving quotitive division problems. For solving partitive division tasks this strategy is far less suitable as information about the number of items per group is not available. Why the children chose the grouping strategy for partitive division cannot be derived from their written answers and/or drawings. Tumusiime and colleagues have raised this question already in their poster presented at PME 43 and called for further research. The study reported in this paper takes up on that and seeks to answer the following research question based on qualitative interviews: How do children who apply a grouping strategy to partitive division tasks find the correct solution, i.e. the number of items per group?

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. 4, pp. 201–208). PME.

THEORETICAL BACKGROUND

When looking for real world related representations of a division problem, e.g. 12 ÷ 3, one numerical task can present two completely differently structured situations depending on the nature of the problem context and can therefore be illustrated in two interpretations of division: partitive and quotitive division. Both represent prototypical mental models of the mathematical division concept, that are fundamental for the understanding and interpretation of real-word related division problems and hence for a comprehensive understanding of division in primary school mathematics. Table 1illustrates the partitive and the quotitive interpretation of the task 12 ÷ 3:

Interpretation of division	Total number of items	Number of groups	Number of items per group
partitive division Twelve children want to split in three equal groups. How many children are in one group?	12 children	3 equal groups	***Wanted*** How many children in each group?
quotitive division Twelve children want to split in groups of three. How many groups of three can they make?	12 children	***Wanted*** How many groups of three?	groups of 3

Table 1: Partitive and quotitive division tasks for 12 ÷ 3

When comparing the two interpretations of division the total number of items that need sharing is the same (12), while they vary in terms of what is wanted: For partitive division the number of groups (three equal groups) is known and the number of items per group is wanted (How many children in each group?). For quotitive division in contrast, the number of items per group is known (groups of three) while the number of groups is wanted (How many groups of three?). Depending on the problem context the two interpretations vary according to the concrete actions as well as with respect to the mathematical structure. Children's solution strategies depend on their individual interpretation of the task as either partitive or quotitive division. For partitive division the main strategy is *sharing one-by-one,* i.e. allocating the items one by one to their specific places until the dividend is exceeded (e.g. see Kouba, 1989). In their drawings the children draw connecting lines between the single objects and their allocations (Wiggelinghoff, 2022). Some children also use estimation strategies. For the strategy *estimate and share,* identified by Axmann & Bönig (1994), the number of items per group is estimated, subsequently the remainders are shared. For quotitive division children superficially use a *grouping strategy*. They organise the items into equal groups representing the number of items per group given in the problem until the dividend is exceeded (Kouba, 1989). The children's respective drawings show a *grouping by circling* (Wiggelinghoff, 2022). In addition, Mulligan (1992) identified three developmental levels of solution strategies for division problems. The strategies vary according to the level of abstractness (increasing) and the level of modelling (decreasing). Level 1 comprises strategies based on direct modelling and counting, using counters or fingers. That includes *sharing one-by-one, estimate and share* as well as *grouping*. Level 2 includes strategies based on counting, addition and subtraction without direct modelling. Frequently these strategies are similar to the ones in level 1, but the children manage to describe their approach and their visualization of the problem. Level 3 includes strategies based on known or derived facts (e.g. addition facts). As the level increases, the strategies used become more elaborate and can be applied more independently of the underlying interpretation of division.

RESEARCH BACKGROUND

This paper is embedded in an international research project (participating countries are Australia, Chile, China, Germany and Uganda) that aims to explore young children's ideas and understanding of early division prior to school instruction. In this project, Grade 1 and Grade 2 students complete a paper and pencil test comprising up to six division problems set in a real-world context prior to formal classroom instruction:

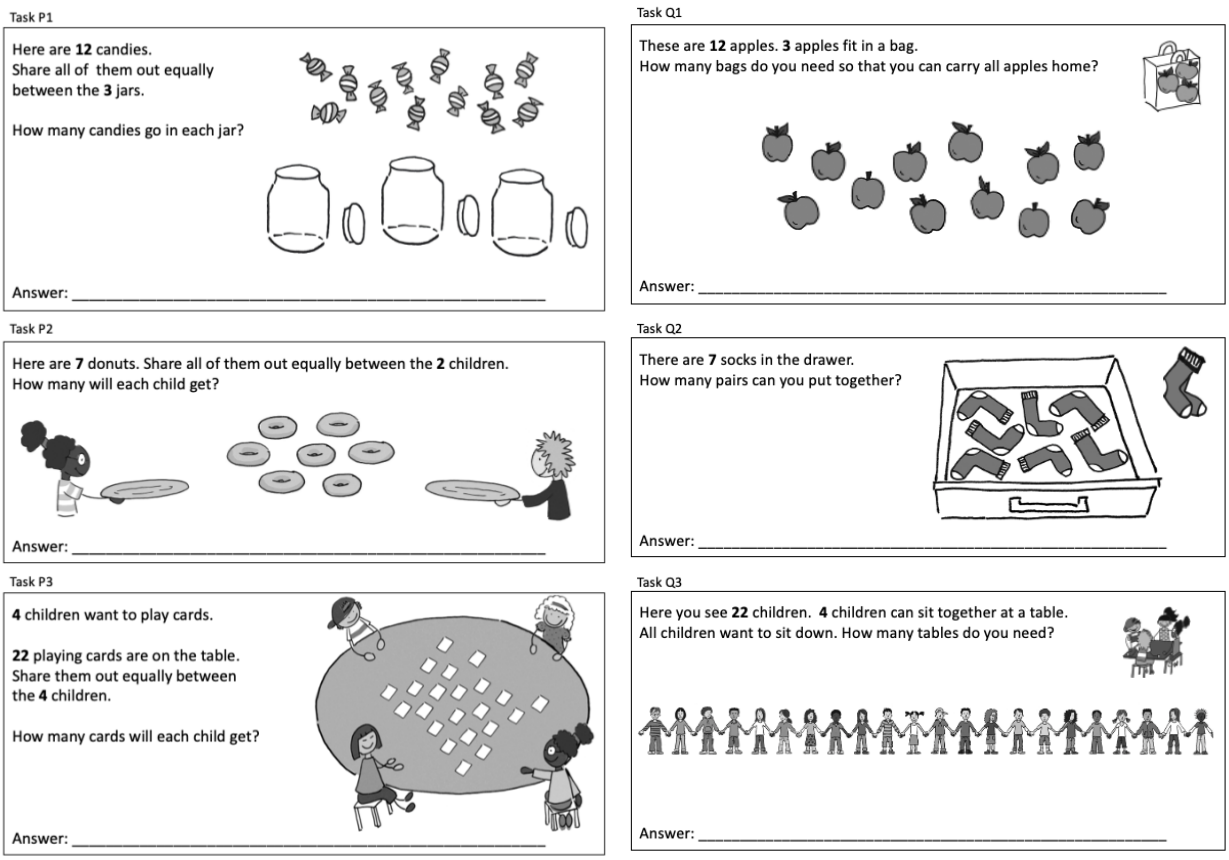


Fig. 1: Partitive and quotitive division tasks with increasing degree of difficulty

First results of the national studies have already been presented at preceding PME conferences. Tumusiime et al. (2019) conducted a study with 96 Grade 1 and Grade 2 students (5- to 8-year-olds) in Uganda. They were presented with a paper and pencil test containing items P1, P2, Q1 and Q3 (see Fig. 1). These tasks were solved with solution rates between 56 and 77 percent. Neither of the two interpretations, quotitive or partitive division, was solved more successfully. These overall high success rates indicate an awareness of division prior to instruction in many of these children. In addition, the children's diverse drawing solution strategies were examined and described. 27 of the 96 children in their sample used a grouping strategy for item P1 (see Fig. 2). The drawing on the right corresponds to the expected strategy of *sharing one-by-one*. The strategy on the left in contrast corresponds to *grouping by circling*. However, the number of items per group is not included in the task and this strategy is therefore not actually applicable. The number of items per group had to be determined in a different way that the respective drawings do not reveal. Cheeseman et al. (2022) have reported results from their study in Australia. Here, 114 Grade 1 students (5- to 6-year-olds) solved all six items of the paper and pencil test with success rates between 33 and 66 percent. Again, no significant differences between partitioning and quotitioning were found in terms of the success rates. Almost three quarters (74%) of children could provide a correct solution to at least one division problem and thus show some awareness of division prior to instruction. Their results were confirmed by the study underlying this paper (see Fig. 1). 74 first graders (6- to 8-year-olds) solved between 23 and 71 percent of the division problems correctly irrespective of their partitive or quotitive context, while 81 percent could provide a correct solution to at least one division problem and show some awareness of division prior to instruction.

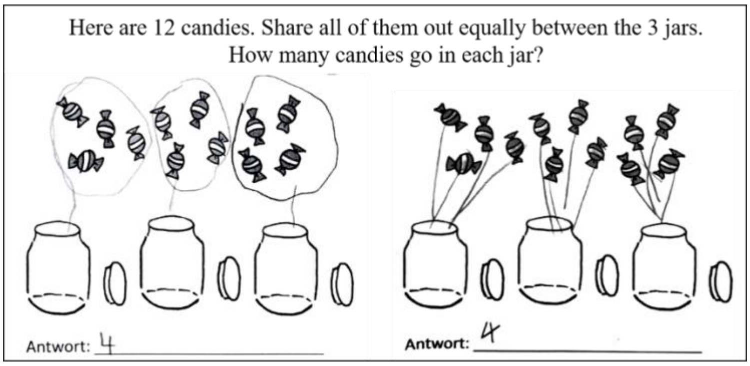


Fig. 2: Grouping (l.) and sharing one-by-one (r.)

METHOD

A sample of five children from the German study took part in the qualitative study reported here. At the time of data collection, they attended a rural elementary school in a small town in western Germany. The children were at the end of their first year of elementary school. Division had not been formally taught in their classes. The children were selected based on their responses to the paper and pencil test described above. All five children had solved at least one of the three partitive division problems P1 to P3 with *grouping by circling*. The method of individual clinical interviews was chosen to reveal the underlying thought processes of the children's solution strategies. For this purpose, the children work on a partitive division problem with numbers identical to task P1, while the context and the presentation of the task are different. The children are given a sheet with 3 mats and 12 red teddies (see Fig. 3) and are asked to solve the following task: “Here are three teddy mats and 12

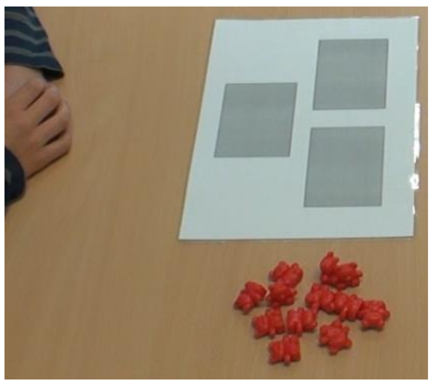


Fig. 3: Partitive division task presented in the interview

teddies. Share the 12 teddies between the three mats, so that there is the same amount of teddies on each mat. How many teddies go on each mat? How did you work that out?” The material encourages an active approach while allowing multiple solution strategies. The children were encouraged to explain their thoughts and procedures when finding the solution. All interviews were video-recorded and transcribed including the actions performed by the children. The transcripts were then analyzed using the qualitative content analysis method (Mayring, 2015). In a first step three main categories were defined deductively based on the interview protocol and the existing literature, i.e. (1) solution strategy and giving the solution, (2) explanation of the solution strategy, and (3) finding alternative solution strategies. All transcript excerpts that clearly show a child's approach and solution strategy as well as the nomination of the result were assigned to the first category. Explanations/ reasonings of children's approaches and solution strategies were assigned to the second category, while alternative solution strategies, which the children were explicitly asked for, were allocated to the third category. In a second step the main categories were differentiated into subcategories based on the interview data, in order to present and evaluate the results in a structured and detailed way. The following example illustrates the coding.

Adam: *takes one teddy at a time and places them in turn on mat 1, mat 2 and mat 3 and then starts again at mat 1, the teddies are therefore shared one by one*

The above excerpt from the transcript refers to Adam’s solution of the “teddies on math task” and is therefore allocated to category 1. From a research point of view, it is also important to analyze which of the numerous strategies was chosen by Adam. These strategies represented in subcategories, i.e. (a) *sharing one-by-one*, (b) *grouping*, (c) *estimation*, (d) *giving the solution without manipulation of materials*. In this example, Adam's approach corresponds to the *sharing one-by-one strategy* and is therefore classified as (sub-)category 1a.

FINDINGS

The qualitative data analyses show three different solution categories and respective thinking and argumentation:

Sharing one by one. Adam (7 years) shares the 12 teddies one by one. This procedure is to be expected for partitive division. For a grouping strategy the number of items per group would have to be known, which Adam makes clear in the argumentation for his approach: “You can't know beforehand that four teddies go on each mat. It is therefore better to share them one by one.” While Adam’s solution does not help contribute to answering the research question, it does illustrate the underlying research gap from a child's line of reasoning.

Estimation and subsequent rearrangement. Two children, Lara and Luke, apply the *estimate and share* strategy. As outlined above, this is also a common procedure for partitive division problems. Lara first places six teddies at once on the first mat and then looks at the remaining teddies for a few seconds. This is followed by a purposeful re-sorting: Two of the six teddies are moved from the first mat to the second mat. The second mat is filled up to four teddies. Lara then places the remaining four teddies on the third, previously empty mat and explains her solution: “I thought there must be six. [...]. But then I looked and tried two. And then I knew I had to do four each.” Lara's first estimate is that six teddies need to go on each mat. She quickly realizes that there are not enough teddies and rearranges two teddies. At this point, she manages to get an overview of the task and realizes that exactly four teddies go on each mat. Luke, on the other hand, takes a slightly different approach using the same strategy. First, he places five teddies at once on the first mat. He then places four teddies on the second mat, leaving three teddies in the pile. After briefly reviewing the situation, a teddy from the first mat is placed on the third mat and the three remaining teddies are added. He explains: “Because I first put five there [points to the first mat]. Then I wanted to put five there [points to the second mat]. Then there were three left [points to the pile], and then I realized ... that there need to be four each.” Luke therefore initially assumes that five teddies go on each mat and recognizes the correct solution in the course if his actions. In comparison to Lara, Luke makes a better estimate, whereas Lara recognizes the correct solution more quickly. Both children provide useful explanations of their solutions in relation to the research question. However, the strategy *estimate and share* does not become obvious in the children’s drawings, as the drawings do not show any corrections that would indicate an initial estimate and subsequent (re)distribution. The explanations of Luke and Lara show that while their drawings in the paper and pencil test look very similar and suggest a *grouping strategy*, their thinking is different from this strategy and also varies slightly between the two children. Furthermore, their explanations reveal important prior knowledge in terms of the understanding of part-whole relationships combined with the ability to make sensible estimates that lead to rapid rearrangements.

Grouping based on preceding calculations**.** Paul and Pia again solve the partitive division problem presented to them in the interview by *grouping*, placing teddies in groups of four successively on the three mats. Both then offer valuable insights as to how they worked out that four teddies go on each mat. Paul argues: “Because four plus four equals eight and eight plus four equals twelve.” Apparently, Paul succeeds in making a connection to known number facts using only the information about the total number of items (12 teddies) and the number of groups (3 mats). He uses repeated addition (4 + 4 + 4 = 12) utilizing an intermediate result (4 + 4 = **8** and **8** + 4 = 12). He obviously has prior knowledge of addition and a part-whole-understanding that he can use to determine the number of items per group for the division task 12 ÷ 3. It is particularly noteworthy that the link between division and other arithmetic skills and prior knowledge is clearly intrinsic in this case. Pia also successfully solves the division problem using a grouping strategy. She states the correct result before she places groups of four teddies successively on the three mats: “Four! Because I know that six plus six is twelve. And four plus 2 ist six and two plus two is four. And then I just put

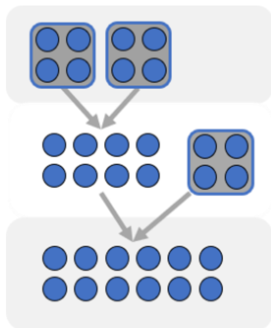


Fig. 4: Paul’s calculation process

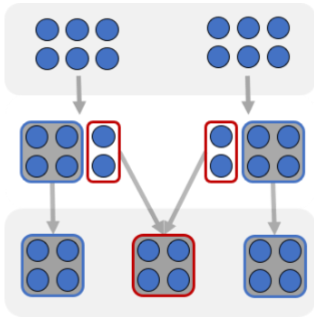


Fig. 5: Pia's calculation process

the two together and that made four." She links the information about the total number of items (12 teddies) with a known number fact: 6 + 6 = 12. Since three equal groups are required, she re-groups 12 into three groups of four. Based on her part-whole understanding she can mentally split both sixes into four and two and then adding the two twos to make four. By decomposing and adding, she manages to re-group two sixes into three fours, which is the correct solution to the task. Again, the link between division and other arithmetic skills and prior knowledge is clearly intrinsic. While Paul's and Pia's approaches are based on fundamentally different ways of thinking, there are obvious parallels: Both use prior knowledge for addition in the form of known number facts as well as their part-whole understanding, indicating an intrinsic (mental) link between division and other arithmetic skills. According to Mulligan (1992), the strategies 1 (sharing one by one) and 2 (estimation and subsequent rearrangement) can be assigned to the lowest level, as direct modeling and solution strategies in the sense of the interpretation of division take place. Paul and Pia, on the other hand, are at least at level 2, as they find solutions based on known number facts and/or number decompositions without having to model their solution process with the teddies. Their strategies are therefore applicable regardless of the underlying interpretation of division and underline Paul's and Pia's deeper understanding of division in this respect.

DISCUSSION AND CONCLUSION

The qualitative sub-study provides relevant findings with respect to the research question. While it was previously not clear how children determine the number of items per group when they use a *grouping* strategy in partitive division problems, the qualitative data analyses revealed two pertinent strategies – *estimation and subsequent rearrangement* and *grouping based on preceding calculations*.

Two children in the sample apparently used estimation strategies involving their part-whole understanding and orientation in number space. This enables them to make solid estimates followed by quick mental rearrangements.

Two other children used calculation strategies such as repeated addition and re-grouping based on previous knowledge of addition and part-whole understanding, enabling them to find the solution without having to model their solution with concrete material. Particularly important for this approach are known number facts and number decompositions. According to Mulligan (1992), such approaches are much more elaborate and can therefore be used by the children independently of the underlying interpretation of division. It is noteworthy that the children use prior arithmetic knowledge to solve division problems without any prompting. This suggests that at least some young children already have intrinsic links between division and other arithmetic skills that can be activated.

When linking this realization to current curriculum documents for early primary school, it challenges the current approach to focus on teaching addition and subtraction (typically taught in Grade 1) a long time before focusing on multiplication and division (frequently taught in Grade 2). Especially when interconnections are intrinsic, it is well worth building on them systematically in school mathematics. The study by Bicknell et al. (2016) has already shown that teaching division from Year 1 to some extent can have a positive impact on the development of operational understanding in general. Furthermore, in a study of children struggling with their mathematics learning Moser Opitz (2007) found that divisional understanding remains a hurdle for many children in learning mathematics right through to secondary school, which is another reason why changes are being called for in the teaching of division in elementary school.

However, the sample of this study is rather small and certainly not generalizable. In order to gain a better understanding of the extent and depth of children's prior to school knowledge and understanding of division, further research into children's intrinsic strategies with a larger sample is clearly needed. It would be helpful to have data from different countries with different curricula and teaching approaches to be able to further discuss necessary changes to classroom instruction.

References

Axmann, A. & Bönig, D. (1994). „15 durch 3, da muß man ja nachdenken, wieviel Dreien in 15 drin sind!". *Grundschulunterricht, 41*(2), 22-24.

Bicknell, B., Young-Loveridge, J. & Nguyen, N. (2016). A design study to develop young children's understanding of multiplication and division. *Mathematics Education Research Journal, 28*(4), 567-583.

Cheeseman, J., Downton, A. & Roche, A. (2022). Ideas of early division prior to formal instruction. In C. Fernández, S. Llinares, A. Gutiérrez, & N. Planas (Eds.), *Proc. 45th Conf. of the Int. Group for the Psychology of Mathematics Educ.* (Vol. 2, pp. 131-138). PME.

Kouba, V. (1989). Children's solution strategies for equivalent set multiplication and division word problems. *Journal for Research in Mathematics Education, 20*(2), 147-158.

Mayring, P. (2015). Qualitative content analysis: Theoretical background and procedures. In Bikner-Ahsbahs, A., Knipping, C., & Presmeg, N. (Eds.), Approaches to qualitative research in mathematics education (pp. 365-380). New York: Springer.

Moser Opitz, E. (2007). *Rechenschwäche/Dyskalkulie.* Bern (CH): Haupt.

Mulligan, J. (1992). Children's solutions to multiplication and division word problems: a longitudinal study. *Mathematics Education Research Journal, 4*(1), 24-41.

Tumusiime, P., Streit-Lehmann, J. & Peter-Koop, A. (2019). *Dividing equally – results of a pilot study on preconceptions of grade 1 to 2 children in Uganda.* Poster presentation at PME 43.

Wiggelinghoff, L. (2022). *Informelles Divisionsverständnis – Eine explorative Studie mit Kindern des ersten Jahrgangs einer Grundschule des gemeinsamen Lernens.* Master Thesis, Bielefeld University, Faculty of Mathematics.

PERFORMANCE OF JUNIOR HIGH SCHOOL STUDENTS' COMPUTATIONAL THINKING IN MATHEMATICAL PROCESS

Lan-Ting Wu and Feng-Jui Hsieh

National Taiwan Normal University

This study explores the performance of junior high school students in computational thinking within mathematical tasks, which were systematically designed based on 4 computational thinking elements and 3 PISA mathematical processes. We employed inductive analysis to explore types of responses from 60 junior high school students, with 30 students from each of the 7th and 9th grades. The results showed that students performed well in decomposition and pattern recognition, but performed relatively weaker in abstraction. Their algorithm designs could be classified into three major types: graph-oriented, direct code-oriented, and pattern code-oriented. The 9th-graders outperformed 7th-graders in algorithmic design. As long as students could design algorithms for simple cases, they had no difficulty with more complex cases.

INTRODUCTION

The rapid development of technology has profoundly influenced our living practices and ways of thinking. In light of this trend, education must be adequately prepared. Scholars have outlined skills deemed essential for the 21st century (Voogt & Pareja, 2010; Wagner, 2014; OECD, 2021). After Wing (2006) coined the term "computational thinking" (CT), scholars also suggest incorporating CT, which includes decomposition, pattern recognition, abstraction, and algorithm design, in school curriculum (Weintrop et al., 2016; Shute, Sun, & Asbell-Clarke, 2017). The latest PISA 2022 mathematics framework (OECD, 2021) recognizes that students should possess not only mathematical thinking (MT) but also CT skills.

Some studies consider problem-solving as one of the commonalities between MT and CT (CSTA, 2011; Rambally, 2015). While CT originates from the field of information technology and often involves using programming to cultivate CT, it doesn't necessarily have to rely on computers (Kaufmann & Stenseth, 2020). Tasks involving CT are classified as either plugged or unplugged, depending on whether computers are used, while MT can be viewed as tasks for problem-solving within specific mathematical content that do not necessarily require the use of computers (Wu & Yang, 2020). It is crucial for teachers to recognize MT within CT and ensure a simultaneous focus on both in planned activities; similarly, the reverse is also essential.

Understanding students' current computational thinking skills in mathematical can greatly assist teachers in developing CT activities in the field of mathematics in the future. Therefore, this study focuses on the following research questions:

1. How do junior high school students perform on CT?

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. 4, pp. 209–216). PME.

2. What are the commonalities and differences of the characteristics exhibited by 7th-graders and 9th-graders in CT?

RESEARCH METHOD

Research framework and instrument

The conceptual framework for understanding students' CT skills in mathematical in this study included two dimensions which were the 4 elements of CT (decomposition, pattern recognition, abstraction, and algorithm design), and the three mathematical processes in mathematical reasoning problems (formulate, interpret, and employ) in the 2022 PISA mathematical literacy framework (OECD, 2021).

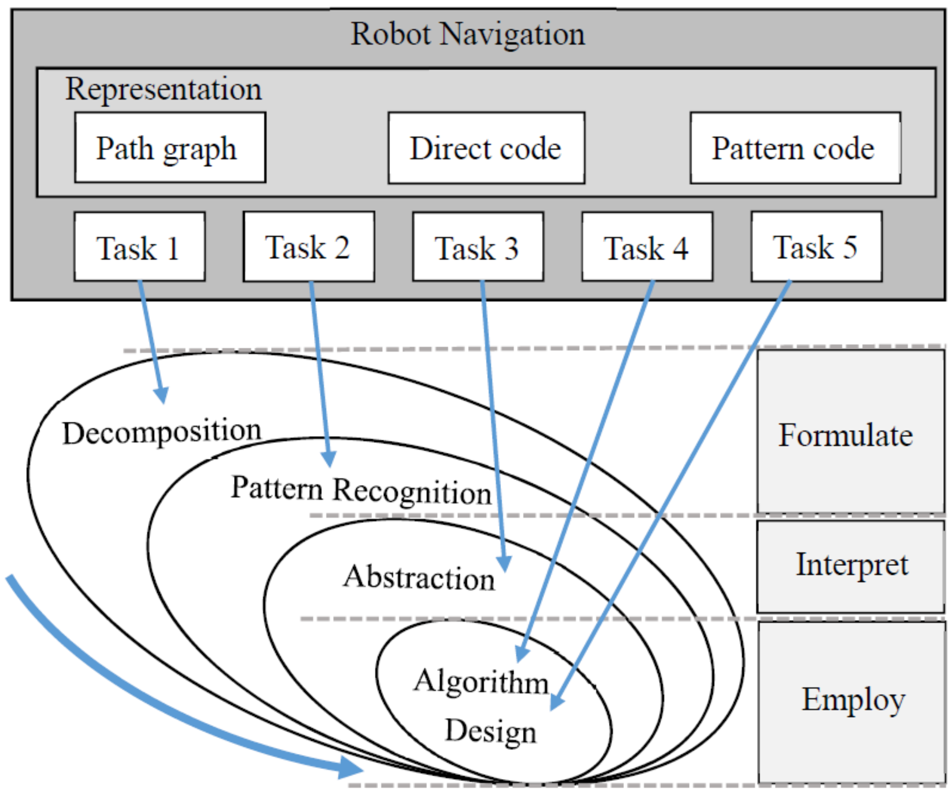


Figure 1: The framework of this study

The instrument used in this study included a series of 5 unplugged tasks, each task corresponding to the two-dimension framework, as shown in Figure 1. In the dimension of mathematical reasoning, each task corresponds to a mathematical process; in the dimension of CT, each task adds an additional CT element to the previous task. Task 1 corresponds to decomposition, and Task 2 corresponds to decomposition and pattern recognition. Both tasks simultaneously correspond to the formulate process. Task 3 adds the abstraction element of CT and corresponds to the interpret process. While the final tasks, Task 4 and 5, both add the algorithm design element and correspond to the employ process.

The tasks used are not explicitly included in our mathematics curriculum, rather inspired by the PISA released items. The tasks were developed through 8 sections of focus group discussion with a math education professor with 40 years of teaching experience, a PhD student with 16 years of teaching, and three master students.

The main theme of the tasks is to manipulate the path of a robot, which can only move upwards or to the right. Students are required to identify, judge, or produce paths that intersect the diagonals of specified m × n grids the most or least times. Task 1 and task 2 guide students to start with a 2 × 1 and a 3 × 2 grid respectively. Students have to identify and create different patterns of paths to demonstrate their ability in “the decomposition, pattern recognition x formulate” dimensions. In Task 3, a path in a 2 × 9 grid is used to introduce two ways of recording the paths, *direct code* and *pattern code* (see Figure 2). Students have to then judge and interpret whether and why the given path matches another *pattern code* (see Figure 3).

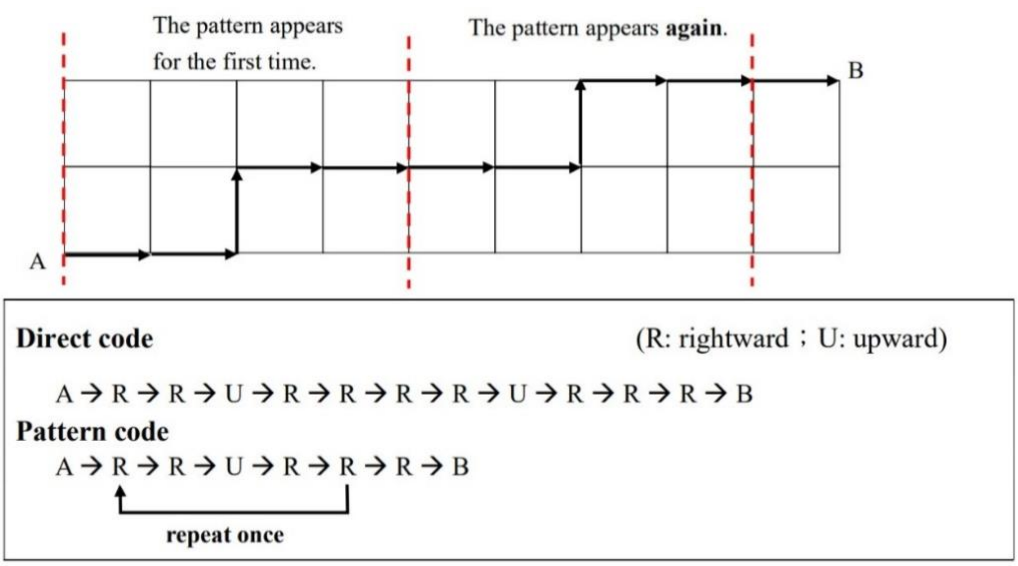


Figure 2: Two methods for recording the robot's path

Pattern code

A → R → R → U → R → R → R → B

repeat once

☐ Correct

☐ Incorrect Why ? ______________________________

Figure 3: An alternative correct *pattern code* to task 3

Task 4 involves a 3×4 grid. Students are asked to determine whether moving upward-first or rightward-first can result in the most intersections with the diagonal. Additionally, students have to produce a path with no intersections with the diagonal and use *pattern code* to record their paths. Task 5 involves a complex 14×15 grid (see figure 4). Students are required to move along the 8x6 grid to point C and then proceeding along the 6x9 grid to point B, and record their paths using the way of *pattern code*.

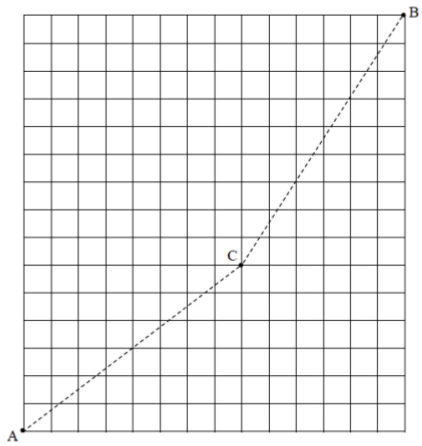


Figure 4: The grid in task 5.

Participants

The sample included 60 junior high school students in Taiwan, with 30 seventh graders and 30 ninth graders from the same school. Different grades were chosen for comparison. The average academic achievements of the students in this school are moderate below national average (in terms of national senior high school entrance examination). Approximately 36% (national average 26%) of the students at this school are identified as needing improvement in mathematics, while 15% (national average 25%) of the students have achieved proficiency in mathematics. These students had not studied similar topics in their math classes previously.

Data collection and analysis

This study collected students' responses on worksheets, comprising 7 pages with a total of 5 tasks. The analysis of students' responses utilized both content analysis and inductive analysis, with the goal of exploring the intricate connections between students' answers, their reasoning processes, and the various elements of CT. Regarding correctness and completeness of students’ responses, we assigned codes of 1 point (complete and correct), and 0 point (incorrect) to examine potential differences in CT skills among students of different grade levels.

RESEARCH FINDINGS

The computational thinking performance of students from Task 1 to Task 5

In this section, all the reported correct percentages were from the last question of each task and represented the final correctness status of that task. Students' performances in the five tasks dropped significantly starting from Task 3. In the formulate process (Task 1 and Task 2), students focused on exploring paths within the smaller grids, observing whether it was more advantageous to move to the right or upward first. Due to the limited number of possible paths and the perceived affordance (Norman, 1999) of the worksheets for students to explore practically, the percent corrects were high. Task 3 required students to determine whether the alternatively given *pattern code* matched the given paths which had originally been recorded using another *pattern code*. Many students believed that the paths should not be the same by simply switching the place of repetition portion of *pattern code*, which resulted in the percent correct much lower than those of the first two tasks (see Table 1).

Percent correct	Task1	Task2	Task3	Task4	Task5
Total (N=60)	77%	81%	57%	48%* (40%)	50%* (25%)

Note. * represents % of *pattern code*. The % of using *direct code* is in parentheses.

Table 1: Percent corrects for Tasks 1 to 5

Task 4 required the paths not to intersect with the diagonal lines, and the constraints were relatively few, allowing for many possible answers. Students primarily chose to repeat two steps, and their response types could be categorized into three major types. Type 1: Moved directly upward and then to the right (22%). Type 2: Moved upward twice, then to the right, and repeated the sequence of moving upward and to the right twice (8%). Type 3: Moved upward twice, then to the right, and repeated the sequence one more time, then moved to the right (17%) (see Figure 5).

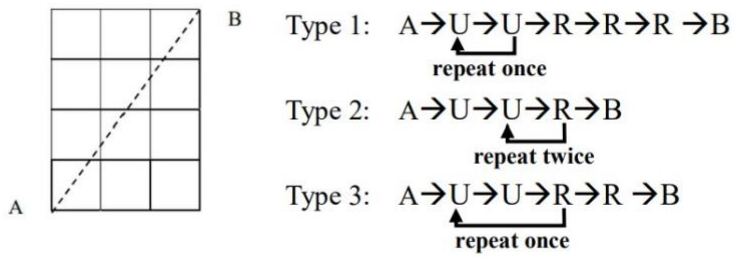


Figure 5: The students' major types of *pattern codes for* Task 4

Comparing the two algorithms (*pattern codes*) listed in Task 3 (Figure 2 and Figure 3) with the algorithm designed by students in Task 4, it shows that students perform better in capturing smaller repeating units during the process of pattern recognition and abstraction. This phenomenon is also evidenced in Task 5.

From students' responses in Task 5, it showed that when facing a large grid, students returned to the stages of decomposition and pattern recognition to observe and explore. After abstraction, they then formulated their algorithm. In this study, students' algorithm designs could be categorized into three major types. The first type, named graph-oriented type (12%), involved recognizing small, repeated patterns from the drawn paths. The algorithms in this category featured short repeating sequences (two steps), sometimes in a right-up loop and other times in an up-right loop. The second type, named *direct code*-oriented type (7%), first transformed the path graph into a *direct code* and then identified only one specific loop (such as a right-up loop) from the *direct code*. As a result, the repeated patterns in the algorithm designs were all the same. The third type, named *pattern code*-oriented type (23%), identified repeated patterns in sub-grids and built the *pattern code* for the whole grid using patterns of the sub-grids. This type of algorithm design had a larger number of steps in a loop; steps in different loops may differ and no steps were left outside the loops (see Figure 6 and see Figure 7 for an example). However, 25% of students still struggled to generate any *pattern codes*, and instead chose *direct code* to record paths.

graph-oriented: A→R→U→R→R→U→U→R→U→U→R→U→B

repeat twice   repeat 5 times   repeat once   repeat once

direct code-oriented: A→R→U→R→R→U→U→R→U→U→R→U→B

repeat twice   repeat 5 times   repeat once   repeat once

pattern code-oriented: A→R→U→R→U→R→U→R→U→R→U→R→U→B

repeat once   repeat twice

Figure 6: The three major types of students' answers to Task 5

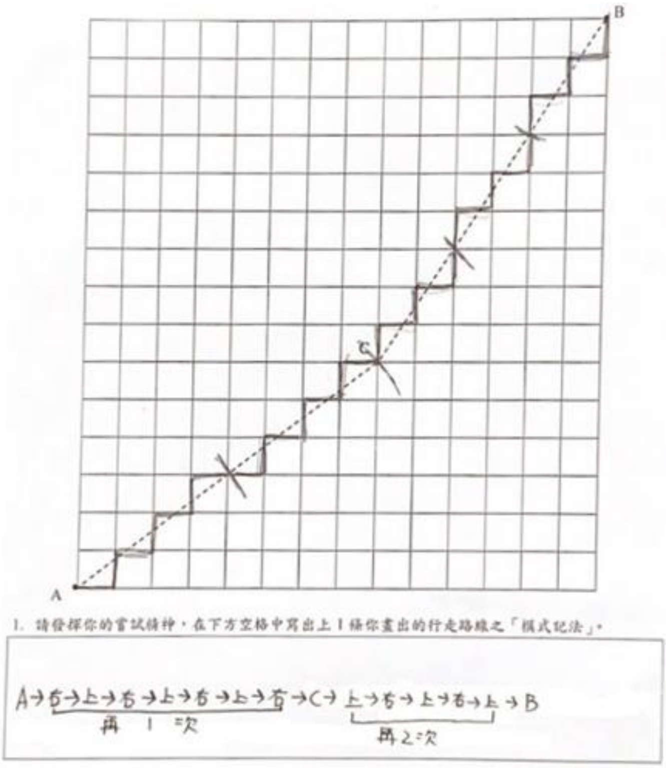


Figure 7: An example of *pattern code*-oriented answer in Task 5

Comparing the percent corrects for Task 4 (48%) and Task 5 (50%), the findings suggest that while students can generate *pattern codes* for moving steps for small grids, they have no difficulty generating *pattern codes* for much larger grids even though being given more complex tasks.

The commonalities and differences in the computational thinking characteristics exhibited by 7th-graders and 9th-graders

The result showed no significant differences in the correctness of answers between 7th-graders and 9th-graders in the first three tasks. In Task 1 and Task 2, which involved decomposition and pattern recognition, the both percent corrects for 7th-graders were

78%, while the percent corrects for 9th-graders were 80% and 83%, respectively. In Task 3, in which abstraction was introduced, although 9th-graders had a higher percent correct (63%) than 7th-graders (50%), the percentages were not significant different (see Table 2).

However, significant differences of percent correct between 7th-graders and 9th-graders were observed in Task 4 ($p < .05$) and Task 5 ($p < .05$), when algorithm design (*pattern code*) was a required element in the measurement. Compared to 7th-graders, 9th-graders could more comfortably use *pattern code* to fulfil task requirements. In Task 4, about the same percentages (85% and 86%) of students in both grades could provide path descriptions. However, only 30% of 7th-graders utilized *pattern code*, compared to 63% of 9th-graders. As shown in Table 2, there were similar results in Task 5. This indicated that 9th-graders outperformed 7th-graders in algorithm design ability.

Percent correct	Task1	Task2	Task3	Task4	Task5
Grade 7 (N=30)	78%	78%	50%	30%* (55%)	30%* (40%)
Grade 9 (N=30)	80%	83%	63%	63%* (23%)	66%* (10%)

Note. * represents % of *pattern code*. The % of using *direct code* is in parentheses.

Table 2: Percent corrects for tasks varied between 7th-grade and 9th-grade students

CONCLUSION

Currently, activities designed to cultivate CT skills often focus on elementary school students' Scratch, programming or STEM (Weintrop et al., 2016). The 2022 PISA, assessing 15-year-old students, explicitly states that “students should possess and be able to demonstrate computational thinking skills as they apply to mathematics as part of their problem-solving practice”. (OECD, 2021). This means the integration of CT in junior high school mathematics classrooms is essential. Based on this claim, this study aims to develop and measure CT skills in junior high schools through mathematical tasks that can identify mathematical processes. We adopted an unplugged approach because it is the dominant approach in Taiwanese mathematics classes.

Although our samples’ mathematical achievement were below Taiwan’s average, they demonstrate high levels of CT elements of decomposition and pattern recognition. This may result from the fact that our tasks that testing these two elements only used small grids, and Taiwanese students are accustomed to approaching difficult tasks through understanding simpler examples, which is an approach often used in math problems in Taiwan’s national senior high school entrance examination. This phenomenon can be seen for the most complex Task 5, where about half of the students can decompose and identify patterns in small units.

Due to space constraints, this paper does not delve into specific features of students' pattern recognition. However, a more comprehensive explanation of students' pattern recognition characteristics will be provided in the conference presentation. And given that students' performance dropped significantly in both grade levels starting from Task 3, which specifies abstraction in the interpret process, future classroom instruction could benefit from placing greater emphasis on the abstraction element of CT and the interpret process. More research is needed to explore strategies for enhancing students' abstraction skills.

References

Computer Science Teachers Association. (2011). Operational Definition of Computational Thinking for K-12 Education. Retrieved from https://csta.acm.org/Curriculum/sub/CurrFiles/CompThinkingFlyer.pdf

Kaufmann, O. T., & Stenseth, B. (2020). Programming in mathematics education. *International Journal of Mathematical Education in Science and Technology*, *52*(7), 1029–1048. doi:10.1080/0020739X.2020.1736349.

Norman, D. A. (1999). Affordances, conventions and design, *Interactions, 6*(3), 38-43. ACM Press.

OECD. (2021). PISA 2021 Mathematics Framework (Draft). OECD Publishing. https://www.oecd.org/pisa/sitedocument/PISA-2021-mathematics-framework.pdf

Rambally, G. (2015). The Synergism of Mathematical Thinking and Computational Thinking, In D. Polly (Ed.), *Cases on Technology Integration in Mathematics Education*, 416-437.

Shute, V. J., Sun, C., & Asbell-Clarke, J. (2017). Demystifying computational thinking. *Educational Research Review, 22*, 142–158. https://doi.org/10.1016/j.edurev.2017.09.003.

Voogt, J., & Pareja, R. N.(2010). *21st century skills*. Enschede, the Netherlands: Universiteit Twente.

Wagner, T. (2014). *The Global Achievement Gap: Updated Edition*. New York, NY: Perseus Books Group.

Wu, W.-R., & Yang, K.-L. (2022). The relationships between computational and mathematical thinking: A review study on tasks, *Cogent Education*, *9*(1). doi: 10.1080/2331186X.2022.2098929.

Weintrop, D., Beheshti, E., Horn, M., Orton, K., Jona, K., Trouille, L., et al., (2016). Defining computational thinking for mathematics and science classrooms. *Journal of Science Education and Technology 25* (1), *127–147*. https://doi.org/10.1007/s10956-015-9581-5.

Wing, J.M. (2006) Computational Thinking. *Communications of the ACM, 49, 33-35*.

AN INVESTIGATION ON THE MATHEMATICAL CREATIVITY OF REGULAR JUNIOR HIGH SCHOOL STUDENTS IN TAIWAN

Yuan Jung Wu and Feng-Jui Hsieh

National Taiwan Normal University

This study explores eighth-grade students' creative thinking skills in the three mathematical processes: formulate, employ, and interpret mathematics in PISA's mathematical literacy framework. Questionnaires capturing the fluency, flexibility, and originality indicators of creative thinking were developed and randomly distributed to 225 students in four regular schools in Taiwan. Inductive analysis was conducted to gain several categories with each several patterns of the responses. Coding rubric was developed. The results show that Taiwanese students performed well in fluency and flexibility in the formulate process but showed only moderate flexibility in the employ process. The research team was surprised as many students' creative answers surpassed expectations.

The 2022 Programme for International Student Assessment (PISA), organized by the OECD, for the first time incorporated 21st century skills, including creativity, in the mathematics test items (OECD, 2018), primarily because creativity is considered crucial for future economic growth and societal development (Binkley et al., 2012) In a cross-country standard and textbook analysis study, supported by the OECD Future of Education and Skills 2030 project, both the national curriculum standards and the higher-order exercises were coded as to whether they included 21st century competencies, including creativity (Schmidt et al. 2022). Inspired by these projects, this study aims to investigate the capacity of regular students' creative thinking in the three mathematical processes: formulate, employ, and interpret, classified by PISA. The purpose of this study is to explore:

(1) the creative thinking products of students in the formulate process.

(2) the creative thinking products of students in the employ process.

(3) the creative thinking products of students in the interpret process.

LITERATURE REVIEW

Creativity has been regarded as an indispensable component in 21st century skills. Binkley et al. (2012) organized the ten skills they have identified into four groupings, the first item in the first grouping "ways of thinking" is creativity and innovation. Many organizations also view creativity as a core 21st century skills (Partnership for 21st century skills, 2012; NRC, 2012). It can be seen from this that experts have restructured their views of creativity, and think about it more as a core skill to be developed for all students rather than as a personality trait exclusive to gifted students (Sternberg, Kaufman & Grigorenko, 2008). Cropley (1992) points out that in terms of

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. 4, pp. 217–224). PME.

teaching students in schools, the nature of creativity refers to a special kind of thinking, often called divergent thinking rather than the other nature of generating novel or creative products. Cropley claims that creativity is "the capacity to get ideas, especially original, inventive and novel ideas" in teaching. Creativity in this situation is also called creative thinking.

Researchers considered creative thinking as an important element in mathematics classrooms for it is an essential thing to support mathematical thinking and communication (Novita & Putra, 2016) and it provides students the opportunity to appreciate the beauty of mathematics (Mann, 2006).

Most researchers believe that mathematical creative thinking is a multi-faceted construct, involving both divergent thinking and convergent thinking (Runco, 1993). However, when evaluating the products of mathematical creative thinking, divergent thinking derived from Guilford (1959) is mainly used (Torrance, 1966). A major approach adopted by researchers to evaluate creative thinking is the use of indicators. One major approach is to use three indicators: fluency, flexibility, and originality (Hollands, 1972; Haylock, 1987; Kim et al., 2003; Lee & Seo, 2003); while another major approach is to add an elaboration indicator to the first three indicators.

The authors believe that preliminary research on regular students' mathematical creativity can ignore elaboration to avoid too many students' frustration when answering questions.

When considering creativity in mathematics, scholars and experts agree that giving multiple solutions is a manifestation of creativity (Arıkan, 2017; Leikin, 2009). Leikin and Lev (2013) used a multiple solution method to explore the relationship between creativity and mathematics achievement. Their research results showed that we cannot use mathematics achievement to judge the level of students' mathematical creativity. In the PISA series of assessments, eight 21st century skills are included in the assessment of mathematical literacy for the first time. PISA emphasizes that these 21st century skills are both supportive of and developed through mathematical literacy (OECD, 2018), one of which is creativity. Although PISA advocates the importance of 21st century skills, in this PISA2022, items were not deliberately developed to incorporate or address 21st century skills. Instead, the identified 21st century skills were incorporated in the items. Items in the 2022 PISA mathematics test were assigned to either mathematical reasoning or one of three mathematical processes: formulate, employ, and interpret mathematics.

METHODOLOGY

Conceptual framework of the study

For creative thinking, we adopted the same construct as Leikin (2009), that creative thinking consists of three indicators: fluency, flexibility, and originality. For mathematical processes we adopted PISA 2022 framework, which includes formulate, employ, and interpret mathematics (OECD, 2018). The conceptual framework of this study is shown in Figure 1.

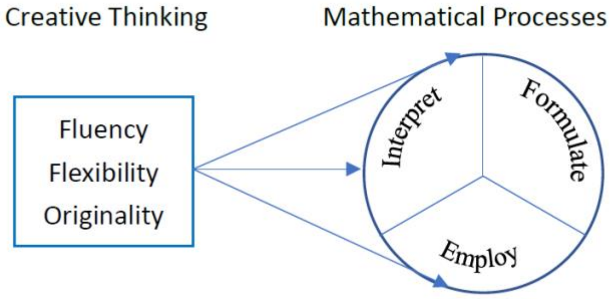


Figure 1: Conceptual framework of the study

Formulate refers to formulating situations mathematically. Employ refers to employing mathematical concepts, facts, procedures and reasoning. Interpret refers to interpreting or evaluating mathematical outcomes. Fluency, flexibility, and originality are classified as generating a variety of ideas, generating different kinds of ideas, and generating unique and novel ideas. (the statistical infrequency of the responses in relation to the peer group, Haylock, 1997).

Design and instrument

This study selected four realistic, authentic, real-world contexts. In each context, nine open-ended items were developed to measure all indicator of creativity in all mathematical processes, for a total of thirty-seven items. The four contexts cover at least four tasks: working on operations, redefinition and reclassification, pattern restructuring, and problem posing. This process was implemented through more than ten focus group discussions, which consisted of up to three professors, two PhDs with an average teaching experience of more than 20 years, and three Master's students. The items were all pilot tested with more than twenty students to test the feasibility of the items and the time required to answer them. Items were remodified accordingly. Based on the pilot study, four questionnaires were developed, each covering only one context with nine to ten items.

In this report, we focus on the results of one context, named "Guessing-Key Game" in which students' task is related to working on operations. The context describes that a class is playing a game related to guessing the real keys while the students only know the public keys obtained by converting the real keys into public keys through a so called "black-box action". The one who develops the most difficult-to-guess "black-box

	Formulate	Employ	Interpret
Fluency	F1a, F1b	E1a, E1b	I1a, I1b
Flexibility	F1a, F1b	E1a, E1b, E2	I1a, I1b
Originality	F2	E3	I2

Table 1: Distribution of items capturing creativity indicators in different math processes

action" wins the game. The ten items, F1a, F2a, …, etc, capture indicators of creativity in different mathematical processes are shown in Table 1.

The prompt for the items specified to formulate process is exemplified as follows: "Please design two 'black-box actions' that are harder to guess than the one designed by Ping. The bigger the difference between your two 'black-box actions' the better." (Ping's design simply adds one to the true keys.)

This study used purposive sampling. Two schools each from the north and south of Taiwan were chosen. The "Guessing-Key Game" questionnaires were randomly distributed to a total of 225 eighth-grade students in these four schools. Students could respond to the questionnaires for 45 minutes. We obtained 222 valid samples.

Data analysis

Six focus group discussions were conducted in two months to generate ways to do inductive analysis and to develop coding rubrics for students' responses. The principle of coding was that each response must be given a response type code, which consisted of several digits. Each digit place represented a category (e.g., the category of *converting ways* and the category of *math ideas used*). Different values in a digit place represented different patterns in that category (e.g., using absolute values and using exponentials are different patterns in the category of math ideas used). The subject's responses were then assessed for fluency according to the number of acceptable responses, flexibility according to the number of different patterns of responses, and originality according to the infrequency patterns of the responses among the samples. Each response was then coded by two coders. The coder reliability for the data reported in this article are all higher than 0.9. For codes where consensus was not reached, coders held in-person or online meetings to consult with another expert in this study to obtain final codes.

RESULTS

Due to space limitations, we only provide some examples and a small portion of the results relating to formulate and employ processes in this report. More results will be given at the conference.

Fluency in the formulate process

Questions capturing fluency in the formulate process were answered by as many as 92% of participants, demonstrating that Taiwanese junior high school students possess a high level of fluency in the formulate process. The most frequently observed patterns involve using first- or second-degree polynomials in the design of "black-box action" (abbreviated as BB action below), accounting for 80% of the cases. The reasons for the high occurrence of these patterns may be that in Taiwan, the mathematics content by the end of grade 8 covers linear functions with rational coefficients, polynomials (mainly 1st and 2nd degree), and their arithmetic actions and factorization. When students provided two different polynomials, they are regarded as equipped with fluency in creativity, but not flexibility.

Flexibility in the formulate process

One criterion for flexibility is when students' two BB actions are classified as having different patterns in any one of the three categories: *converting ways*, *math ideas used*, and *forms*. A total of 60% students meets this criterion. Figure 2 shows examples of responses carrying great variation of patterns in *converting ways* (Student A) and *math ideas used* as well as *forms* (Student B).

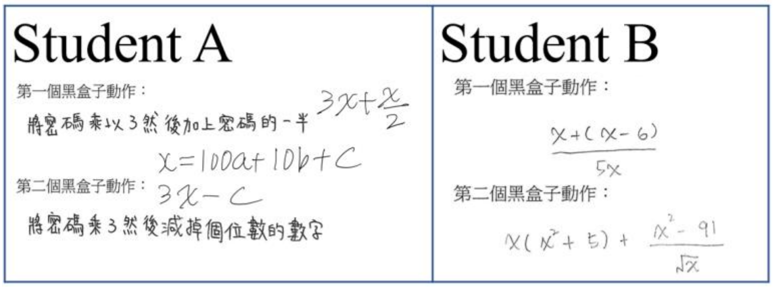


Figure 2: Responses show flexibility in the formulate process.

The first BB actions of student A is of the type $ax + b$. Student A described his/her second BB action as "Multiply the key by 3 and subtract the ones digit from the result". The patterns of the *converting way* switched from "function relations" to "restructuring or operating on place digits". The patterns of *math idea used* for Student B switched from "rational function" to "combination of radical expression, fractions, polynomials, and prime numbers".

Originality in the formulate process

This study classified responses that were particularly unique and difficult for peers to think of (less than 5% occurrences) as possessing originality, specifically, those with non-function patterns or very special function relations in *converting ways* and patterns that linked to other subject or unlearned mathematical domains in the category of *expanding*.

This study classifies as original responses that are particularly unique and difficult for peers to come up with (occurrence less than 5%), specifically those that use non-functional relations or very specific functional relations in the category of *converting ways*, or links to other areas or unstudied math areas in the category of *expanding* (see Figure 3 for examples).

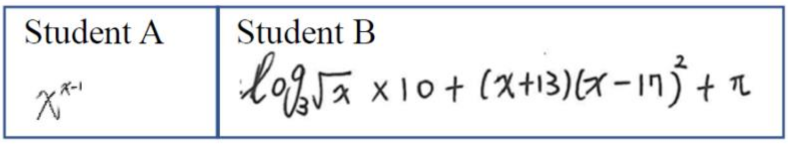


Figure 3: Responses show originality in the formulate process.

Fluency in the employ process

For the employ process, the items provide students with Ping's teacher's true key (10) and public key (88) and require students to guess the BB actions that the teacher might designed to convert 10 to 88, with no restriction (items E1a & E1b) and with the restriction of using squares (E2) and square roots (E3).

A total of 72% of students give two different appropriate BB actions when no restriction is given, showing their fluency capacity. The two BB actions with the highest frequency fall in the type of two-term addition $ax + b$, such as $8x + 8$ and $x + 78$.

Flexibility in the employ process

This study classifies switching patterns in any one of the following categories as equipping flexibility: *expression structure*, *number system involved*, *complexity levels*, and *expending*. Student A shown in Figure 4 switched patterns from $ax + b$ to including absolute value symbol in the category of *expression structure*, while student B switched patterns from $ax + b$ to including arithmetic sequence.

In terms of restricting to inclusion of squares and square roots in BB actions, 69% and 43% of students were able to provide correct answers that differed from their first two BB actions, respectively. These students are considered equipped with flexibility; if further extended to use rational or irrational numbers, they are considered to have higher flexibility.

Student C uses a linear function but puts the square on constants, and uses the sum of squares of consecutive positive integers. This shows that Student C is proficient in number actions and sequence concepts, and meanwhile has flexibility.

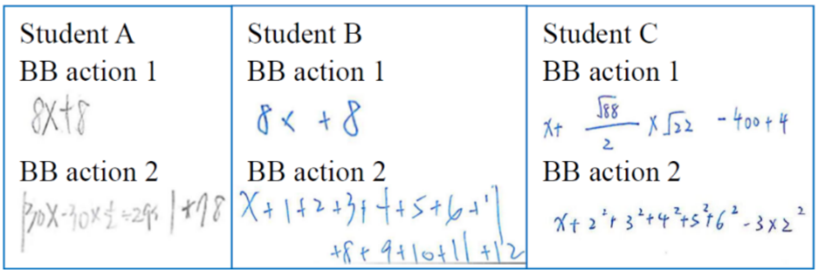


Figure 4: Responses show flexibility in the employ process.

Originality in the employ process

The criterion of less than 5% occurrences is still used to single out originality responses. Two examples of originality response are shown in Figure 5.

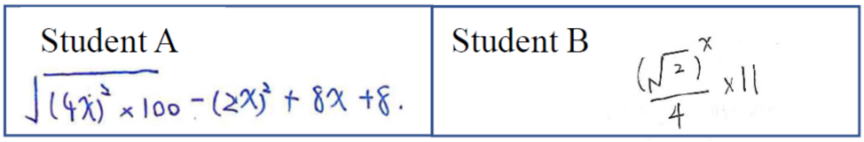


Figure 5: Responses show originality in the employ process.

Student A is able to develop the relationship between squares and square roots $\sqrt{(4x)^2 \times 100} - (2x)^2 = 0$, which is quite a unique idea. Most students' appropriated responses use perfect square numbers and adjust the values with a constant or coefficient, such as $22 \times \sqrt{x+6}$、$\sqrt{x+6}+84$、$9x-\sqrt{4}$ and $\sqrt{x^2+21}-33$. Student B is able to think of 32 as the power of 2, and then think of the relationship between the square root of 10 and employs the concept of exponents.

CONCLUSION

In the context relating to mathematical operation, Taiwanese eighth graders show moderate (about 50%) to a high degree (about 95%) of fluency and flexibility. Students performed better in formulate mathematics than in employ mathematics. It is possible that in the employ process, to meet the situations specified in the items restricts the possibility of free creation. Comparing the results with the results obtained by the other three contexts in the whole study may provide a more precise and evidenced conclusion. In terms of originality, the criterion of less than 5% occurrence patterns is bound to result in only originality in a few students. Other possible ways should be considered to re-examine the criteria for originality. Students' responses show that when constructing actions that convert values, they are not restricted to the learned mathematical concepts or expressions; rational functions, expressions with absolute value, higher order powers or powers with unknowns, combination of a wide range of operations are all possible student responses. The results of this study surprised the research team because many students' creative answers exceeded the team's expectations.

References

Arıkan, E. E. (2017). Is there a relationship between creativity and mathematical creativity? *Journal of Education and Learning*. *6*(4): 239. http://doi.org//10.5539/jel.v6n4p239

Ball, D. L. (1990). Prospective elementary and secondary teachers' understanding of division. *Journal for Research in Mathematics Education*, *21*(2), 132–144.

Binkley, M., Erstad, O, Herman, J , Raizen S., Ripley, M., Miller-Ricci, M., & Rumble M.. (2012). Defining twenty-first century skills. P. Griffin et al. (eds.), Assessment and Teaching of 21st Century Skills, 17.http://doi.org//10.1007/978-94-007-2324-5_2, © Springer Science Business Media B.V. 2012. P.17-66.

Chapman, O. (2003). Facilitating peer interactions in learning mathematics: Teachers' practical knowledge. In M. J. Høines & A. B. Fuglestad (Eds.), *Proc. 28th Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 2, pp. 191–198). PME.

Cropley, A. J. (1992): More ways than one: Fostering creativity. Ablex Publishing Corporation.

Guilford, J. P. (1959). Traits of creativity, in H. Anderson (ed.), Creativity and Its Cultivation. Harper.

Haylock, D. W. (1987). A framework for assessing mathematical creativity in school children. *Educational Studies in Mathematics, 18*, 59-74.

Haylock, D. W. (1997). Recognizing mathematical creativity in school children. *International Reviews on Mathematical Education*, *29*, 68-74.

Hollands, R. (1972). Educational technology: Aims and objectives in teaching mathematics, *Mathematics in School, 1*(6), 22-23.

Kim, H., Cho, S., Ahn, D. (2003). Development of mathematical creative problem solving ability test for identification of gifted in math. *Gifted Education International, 18*, 184-193.

Lee, K. S., & Seo, J. J. (2003). A development of the test for mathematical creative problem solving ability. *Research in Mathematical Education*, *7*(3), 163-189.

Leikin, R. (2009). Exploring mathematical creativity using multiple solution tasks. In R. Leikin, A. Berman, & B. Koichu (Eds.), *Creativity in Mathematics and the Education of Gifted Students* (pp. 129-145). Rotterdam: Sense Publishers.

Leikin, R., & Lev, M. (2013). Mathematical creativity in generally gifted and mathematically excelling adolescents: What makes the difference? *ZDM—The International Journal on Mathematics Education*, *45*(2), 183-197.

Mann, E. L. (2006). Creativity: The Essence of Mathematics. *Journal for the Education of the Gifted*, *30*(2), 236-260. https://doi.org/10.4219/jeg-2006-264.

McDonough, A., & Clarke, D. (2002). Describing the practice of effective teachers of mathematics in the early years. In N. A. Pateman, B. J. Doherty, & J. Zilliox (Eds.), *Proc. 27th Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 3, pp. 261–268). PME.

National Research Council. (2012). *Education for Life and Work: Developing Transferable Knowledge and Skills in the 21st Century.* Committee on Defining Deeper Learning and 21st Century Skills, J.W. Pellegrino and M.L. Hilton (Eds.). Board on Testing and Assessment and Board on Science Education, Division of Behavioural and Social Sciences and Education. Washington, DC: The National Academies Press.

Novita, R., & Putra, M. (2016). Using task like PISA's problem to support student's creativity in mathematics. *Journal on Mathematics Education*. *7*(1), January 2016, pp. 33-44.

OECD (2018). PISA 2021 Mathematics Framework (Draft). https://www.oecd.org/pisa/pisaproducts/pisa-2021-mathematics-framework-draft.pdf

Partnership for 21st century skills. (2012). 21 century skills map. https://files.eric.ed.gov/fulltext/ED543032.pdf

Runco M. (1993). Divergent Thinking, Creativity, and Giftedness. *Gifted Child Quarterly*. 1993;37(1):16-22. https://doi.org/10.1177/001698629303700103

Schmidt, W., et al. (2022), "When practice meets policy in mathematics education: A 19 country/jurisdiction case study", *OECD Education Working Papers*, No. 268, OECD Publishing, Paris, https://doi.org/10.1787/07d0eb7d-en.

Sternberg, R. J., Kaufman, J. C. & Grigorenko, E. L. (2008). Applied Intelligence. Cambridge University Pres.

Torrance, E. P. (1966). Torrance Tests of Creative Thinking: Norms Technical Manual, Personnel Press.

WHAT NOVICE MATHEMATICS TEACHERS PERCEIVED IN ASSESSING STUDENTS' LEARNING OF FUNCTIONS

Runyu Zhang[1], Shuhui Li[1], and Qiaoping Zhang[2]

1. East China Normal University; 2. The Education University of Hong Kong

This study explores the perceptions of 96 novice mathematics teachers on assessing students' learning of functions via a collaborative task of constructing a specific test in China. By analyzing the 23 teacher-constructed tests, the study reveals that these teachers demonstrated strong subject knowledge in designing mathematics tests and tended to construct more high-level questions with an object-level functional thinking focus, multiple steps, and high cognitive loads, aligned to or above the curriculum standards and presented in a purely mathematics context using mixed representations. The results provide evidence to explain the possible gaps between teachers' intended curriculum and attained curriculum, and also reflect the values of novice mathematics teachers in assessing students' learning of functions.

INTRODUCTION

Assessment can effectively serve as a tool to measure students' learning outcomes, as well as to validate the curriculum they have attained (Santos & Cai, 2016). Existing research on mathematics assessment usually focused on the assessment task design (e.g., Demosthenous et al., 2021), the context of assessment (e.g., Zhang et al., 2021), and formative assessment (e.g., Baird, 2010), little is known about how teachers construct tests (Becevic, 2023), as well as what teachers perceive in the assessment.

Teacher-constructed test (TCT), a test created or selected by teachers for assessment (Goos, 2020), is a crucial link between learning objectives and assessment of and for teaching, which can function as a method to coordinate the intended, implemented, and attained curriculum (Becevic, 2023). Teachers largely rely on data from TCT to make decisions about students' knowledge (DiDonato-Barnes et al., 2014). As a component of the attained curriculum, TCT can reflect what teachers value in teaching. However, researchers reported that TCT was often not well-written and expressed concerns about the poor quality of TCT (e.g., Watt, 2005; Wellberg, 2023).

Previous studies showed several challenges for teachers to construct a test (e.g., Wiggins, 1992), and new teachers usually feel not confident in testing (Burke, 2009). Therefore, in this study, we aim to examine what Chinese novice mathematics teachers perceive in assessing students' learning of functions. Specifically, the following research questions are addressed: (1) How many aspects of students' learning of functions are emphasized by novice teachers? (2) What kinds of assessment problems are highlighted by novice teachers? (3) Which principles for assessment design are utilized by novice teachers? (4) To what extent does the assessment align with the mathematics curriculum standards?

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. 4, pp. 225–232). PME.

RELATED THEORETICAL CONCEPTS AND STUDIES

Principles for Mathematics Assessment Design

An effective assessment of students' learning outcomes requires a thorough and precise record of student performance, along with a clear scoring guide to derive criterion scores (Medley, 1987). Pathak (2023) identified significant concerns such as the overall insufficient emphasis on important aspects, poorly crafted questions, and tests that are too predictable and lack clarity. Traditional mathematics tests typically focused on the repetition of learned procedures, even though they were capable of assessing a wide range of mathematical capabilities if set appropriately (Watt, 2005). Drawing from Medley's (1987) work, the first design principle for constructing an assessment is the necessity of *precision and clarity*.

The second principle is related to the *format* of assessment problems. Murphy et al. (2023) recommended teachers use a variety of test formats (e.g., cued recall, multiple-choice, and true/false). A mixture of multiple-choice, fill-in-the-blank, and short-answer questions is standard in high-stakes exams in China, e.g., the College Entrance Examination-Mathematics (CEE-M). As teachers intend to get students familiar with the CEE-M, teacher-constructed tests usually follow similar settings.

The third principle considers the *arrangement* of assessment problems by difficulty, typically sorted as easy-to-hard, hard-to-easy, and random (Brenner, 1964). Research on the arrangement is divided; some indicated that starting with easier problems improves students' performance (Hodson, 1984), while other studies found no significant effect (Plake, 1981) or even adverse experiences (Bieleke et al., 2021). We adopt Brenner's categorization and evaluate problem difficulty based on the competency level outlined in curriculum standards.

Teacher-constructed Tests in Functions

The construction of the function concept tends to be more complicated than expected with regard to students. Researchers reported that many students have difficulties in translating and converting between various representations of functions, manipulating symbols, and thinking of functions as objects (e.g., Elia et al., 2007). Even though function is emphasized in the curriculum standards, how teachers carry out an appropriate assessment in students' learning of functions is underexplored.

Reminded by the students' learning difficulties and teachers' challenges in designing tests in terms of functions, we try to offer a more in-depth analysis on functional thinking aspects teachers emphasized to assess students' learning of functions. Therefore, in this study, we adopt Lichti & Roth's (2019) model of *functional thinking*, which includes three levels: (1) *mapping* (Complete a pair of values or identify a given pair of values using a graph, a table, etc.); (2) *covariation* (Use the absolute change, the slope, or the rate of change; or use only the slope of a graph to solve a task); and (3) *object* (Consider various aspects of a function at the same time to classify and relate it to other representation; or connect different representations as a whole, etc.).

Gaps between Intended Curriculum and Attained Curriculum

Schmidt et al. (2005) argued that assessment must align with curriculum standards, textbooks, and actual school practice. Inspired by this, we evaluate the alignment between the competency level demanded by curriculum standards (knowing, understanding, grasping, and applying) and the tests created by novice teachers. The *curriculum standard alignment* is categorized into three levels: *below standards*, *meeting standards*, and *above standards*.

When assessing student's understanding of school curriculum, textbooks are always considered as one of the important resources for teachers (Mullis et al., 2012). To reflect *textbook-problem relevance*, the present study employs Stein et al.'s (2000) *cognitive load* analysis (memorization, procedures without connections, procedures with connections, and doing mathematics), and three problem types generated from Li's (2020) textbook-problem analysis, including *steps required for a solution* (single-step or multi-step), *representations used* (text or mixed representation: text & table, text & graph, and text & formula), and *context* (purely mathematics or real-life).

In summary, the framework for analyzing teacher-constructed tests in this study comprises four aspects: (1) principles for assessment design (encompasses the format of questions, their arrangements, and the precision and clarity of the questions), (2) functional thinking, (3) curriculum standard alignment, and (4) textbook-problem relevance (evaluates the context of questions, steps required for a solution, representations used, and the level of cognitive load). Unique indicators for each aspect establish the criteria for a comprehensive evaluation of the assessment.

METHODS

This study was part of a project examining novice secondary mathematics teachers' noticing of students' functional thinking. Overall, 96 novice teachers from a normal university in China attending a training course on mathematics assessment participated in this study in August 2023. They voluntarily formed 23 groups: one group with 3 teachers, 17 groups with 4 teachers each, and 5 groups with 5 teachers each. After several sessions on how to evaluate assessment (e.g., difficulty, discrimination, reliability, and validity) and principles for assessment design (covering assessment purposes, problem types, as well as steps and techniques to develop an exam paper), the course instructor assigned an in-class collaborative group task as follows: "Design a test comprising 8-10 questions with total 100 points, focusing on functions."

We collected back 23 teacher-constructed tests with 231 questions (5 tests have more than 10 questions). The majority of the questions covered concepts, properties, and applications of function (64.7%), followed by linear function (13.7%), trigonometric function (10.0%), logarithmic function (5.4%), inverse proportional function (3.3%), power function (2.5%), and exponential function (0.4%). Figure 1 provides an overview of four dimensions we used to reveal novice teachers' perceptions in assessing students' learning of functions.

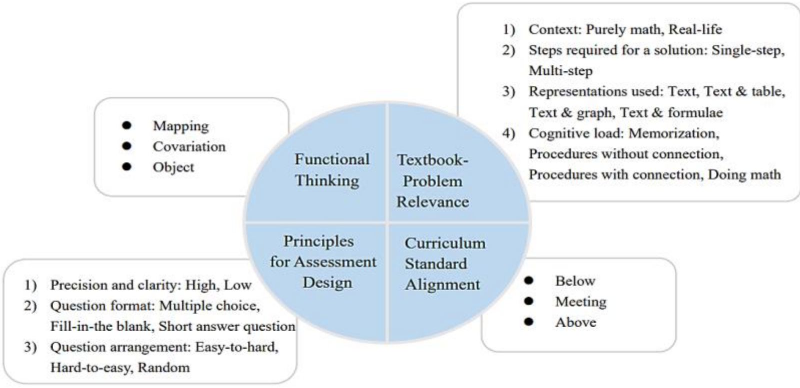


Figure 1: Four dimensions to analyze teacher-constructed tests

For each question, we assigned codes for four dimensions, including functional thinking, textbook-problem relevance (context, steps required for a solution, representations used, and cognitive load), principles for assessment design (precision and clarity, question format, and question arrangement), and curriculum standard alignment. Consider the example question: "Which choice is the correct intersection coordinate between the linear function y=3x+6 and the x-axis? A. (2,0), B. (6,0), C. (-2,0), D. (0,6)." The codes would be: Mapping, Purely mathematics, Single-step, Text & formula, Procedure without connections, High precision and clarity, Multiple-choice, Easy, and Below standards. To verify coding reliability, two researchers independently coded each question, achieving a 93.52% agreement rate. For data analysis, we calculated the quantity and points of questions per exam across nine indicators (see Figure 1) to identify trends in teacher-constructed tests for assessing students' functional thinking.

RESULTS

Functional thinking: This study adopts three functional thinking levels—mapping, covariation, and object—to explore how many aspects of students' learning of functions are emphasized by novice secondary teachers in assessment. Table 1 shows the frequency of questions per test and their corresponding points across three levels.

Level	Number of questions per test	Points of questions per test
Mapping	2.4 (22.9%)	20.1 (20.1%)
Covariation	2.4 (22.9%)	18.2 (18.2%)
Object	5.7 (54.2%)	61.7 (61.7%)

Table 1: Frequency of questions across three functional thinking levels

As shown in Table 1, the majority of the questions belong to object level (54.2%), followed by the same proportion of questions target mapping and covariation level (22.9%; 22.9%). Similar trends exist for the points of questions per test, in which on average 61.7 (a total score of 100) assessing object level, followed by 20.1 for mapping level and 18.2 for covariation level per test.

Textbook-problem relevance: This study analyzes teachers' preferences for assessment problem features: context, representations used, steps required for a solution, and cognitive load. Figure 2 highlights notable differences across these four features. For context, only 13.3% of the questions relate to real-life scenarios, while over 85% are purely mathematical. With regard to representations used, a majority of the questions (75.5%) use a combination of text and formula, while a mere 7.5% are presented solely in text. Mixed text and graph representations account for 16.0%.

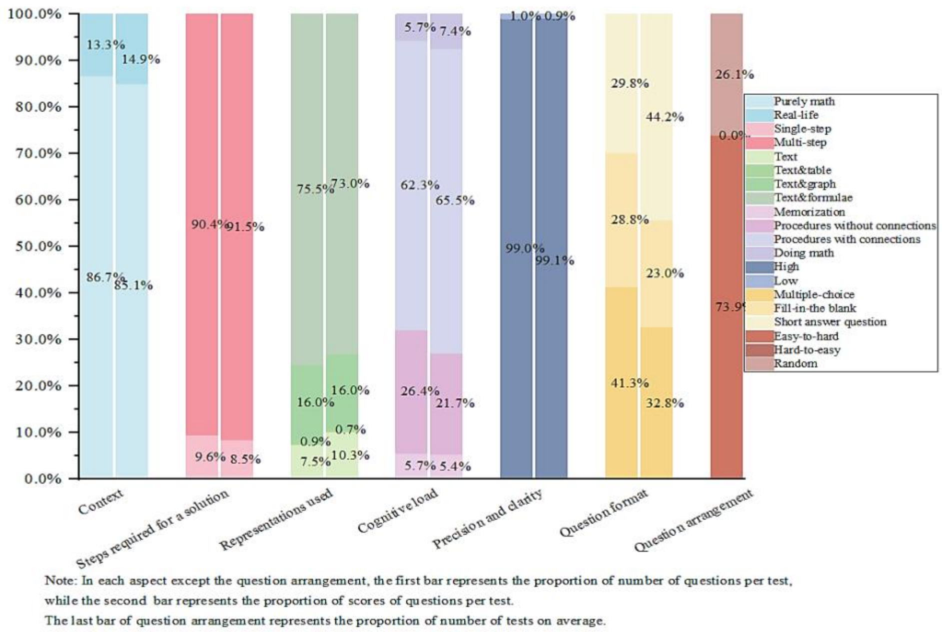


Figure 2: Distributions of questions across textbook-problem relevance & principles for assessment design

For steps required for a solution, the study shows that over 90% of the questions are multi-step, suggesting a preference for more complex problem-solving that invokes higher cognitive demands. This aligns with the cognitive load analysis results, where approximately 68% of the questions are classified as high-level tasks, and a majority (62.3%) involve procedures with mathematical connections.

Principles for assessment design: The study evaluates the design principles employed by teachers by examining precision and clarity, question format, and question arrangement. According to the data presented in Figure 2, a particularly noteworthy outcome is the extreme high level of precision and clarity (99%), signifying that almost all the questions proposed are clearly articulated and mathematically precise.

Regarding the format of questions, inconsistencies between the number of questions and the points assigned to questions per exam are evident in Figure 3. Novice teachers prefer to design a greater number of multiple-choice questions (41.3%) and a comparable number of fill-in-the-blank and short-answer questions (28.8% and 29.8%, respectively), yet they allocate more points to short-answer questions (44.2%). In other words, these teachers use three formats in a relatively even distribution in their teacher-constructed tests but tend to assign more points to short-answer questions. Regarding question arrangement, approximately 73.9% of these teachers design tests with an easy-to-hard arrangement, in contrast to a hard-to-easy or a random sequence which account for 0% and 26.1%, respectively.

Curriculum standard alignment: This study categorizes questions into three levels––below, meeting, and above––to investigate to what extent the teacher-constructed test aligns with the mathematics curriculum standards. Table 2 presents the distribution of questions across three levels.

Level	Number of questions per exam	Points of questions per exam
Below	1.9 (17.8%)	13.9 (13.9%)
Meeting	5.3 (50.5%)	51.8 (51.8%)
Above	3.3 (31.5%)	34.3 (34.3%)

Table 2: Frequency of questions across three alignment levels

From Table 2, about half of the test questions (50.5%) are aligned with the curriculum standards. Consistent with the above findings of a large proportion of high cognitive load questions, 31.5% of the questions are above the standards, and only 17.8% of the total questions are below the standards. Similar trends exist for the assigned points of questions across three levels. It implies that novice teachers in our study tend to design questions above or the same as curriculum standards. In other words, these teachers prefer to construct a challenging test, rather than a standard-aligned one.

Overall, when assessing students' learning of functions, novice teachers tend to construct a test with: (1) More object-level, less but balanced mapping-level and covariation-level questions; (2) More purely mathematics, multi-step, and high cognitive loads questions presented in the text and formula representation; (3) Precise and clear questions arranged in an easy-to-hard sequence, with slightly more multiple-choice, less fill-in-the-blank and short-answer questions while more points assigned to short-answer questions; (4) More questions meet or above the curriculum standards.

DISCUSSION AND CONCLUSION

This study aims to investigate what novice secondary mathematics teachers perceive in constructing a test to assess students' understanding of functions. The most salient finding is that these teachers tend to construct tests that favour high-level questions, such as those that require object-level functional thinking, multi-step, and high cognitive loads, which align with or exceed curriculum standards. These questions are often presented in a mixed representation within a purely mathematical context. This suggests a discrepancy between the intended curriculum and the curriculum that is actually taught and assessed by teachers.

The research further indicates that teachers possess strong subject matter knowledge, as evidenced by the high clarity and mathematical precision of almost all questions in their constructed tests. Additionally, they can utilize some principles for assessment design, such as using different formats of questions and following an easy-to-hard arrangement. Future investigation could be conducted to explore the rationale behind their construction and gain insights on what kind of support or training for in-service teachers (especially novice) when designing good-quality tests.

There are some limitations in research methodology. Instead of counting and comparing the number of different formats of questions, an in-depth analysis can be conducted for representative questions and tests in the future. Profile analysis could also be employed to identify patterns in teacher-constructed tests and determine whether there are significant differences between groups of tests. Furthermore, as topic differs in mathematics teaching and assessment, future studies could focus on more mathematical topics using the theoretical framework proposed in this study, which may help to promote teachers' assessment literacy.

Acknowledgment

This study has been partly funded by the Education University of Hong Kong Departmental Research Grant (Project No. 04863) and research grant from the Ministry of Education of the People's Republic of China (Award No: 22YJC880035).

References

Baird, J. A. (2010). Beliefs and practice in teacher assessment. *Assessment in Education: Principles, Policy & Practice*, *17*(1), 1-5.

Becevic, S. (2023). When teachers construct tests for assessing students' competencies: A taxonomy. *Education Studies in Mathematics*, *114*(2), 315–336.

Bieleke, M., Goetz, T., Krannich, M., Roos, A.-L., & Yanagida, T. (2021). Starting tests with easy versus difficult tasks: Effects on appraisals and emotions. *Journal of Experimental Education*, *91*(2), 317-335.

Brenner, M. H. (1964). Test difficulty, reliability, and discrimination as functions of item difficulty order. *Journal of Applied Psychology*, *48*(2), 98-100.

Burke, K. (2009). *How to assess authentic learning*. Corwin.

Demosthenous, E., Christou, C., & Pitta-Pantazi, D. (2021). Mathematics classroom assessment: A framework for designing assessment tasks and interpreting students' responses. *European Journal of Investigation in Health, Psychology and Education*, *11*(3), 1088-1106.

DiDonato-Barnes, N., Fives, H., & Krause, E. S. (2014). Using a table of specifications to improve teacher-constructed traditional tests: An experimental design. *Assessment in Education: Principles, Policy & Practice*, *21*(1), 90-108.

Elia, I., Panaoura, A., Eracleous, A., & Gagatsis, A. (2007). Relations between secondary pupils' conceptions about functions and problem solving in different representations. *International Journal of Science and Mathematics Education*, *5*(3), 533-556.

Goos, M. (2020). Mathematics classroom assessment. In S. Lerman (Ed.), *Encyclopedia of mathematics education* (pp. 572–576). Springer International Publishing.

Hodson, D. (1984). The effect of changes in item sequence on student performance in a multiple-choice chemistry test. *Journal of Research in Science Teaching, 21*(5), 489–495.

Li, S. (2020). *A comparative study of directional connections in popular US and Chinese high school mathematics textbook problems*. Doctoral dissertation, Columbia University.

Lichti, M., & Roth, J. (2019). Functional thinking—A three-dimensional construct? *Journal Für Mathematik-Didaktik*, *2*(40), 169-195.

Medley, D. M. (1987). Criteria for evaluating teaching. In M. J. Dunkin (Ed.), *The international encyclopedia of teaching and teacher education* (pp. 169–180). Pergamon Press.

Mullis, I. V. S., Martin, M. O., Foy, P., & Arora, A. (2012). *TIMSS 2011 international results in mathematics*. Boston College.

Murphy, D. H., Little, J. L., & Bjork, E. L. (2023). The value of using tests in education as tools for learning—Not just for assessment. *Educational Psychology Review*, *35*(3), 89.

Pathak, S. (2023). Functions and their graphs in high school mathematics in Nepal. In M. Ayalon, B. Koichu, R. Leikin, L. Rubel & M. Tabach (Eds.), *Proc. 46th Conf. of the Int. Group for the Psychology of Mathematics Education* (Vol. 1, pp. 304). PME.

Plake, B. S., Thompson, P. A., & Lowry, S. R. (1981). Effects of item arrangement, knowledge of arrangement and test anxiety on two scoring methods. *Journal of Experimental Education*, *49*(4), 214–219.

Santos, L., & Cai, J. (2016). Curriculum and assessment. In A. Gutiérrez, G. Leder, & P. Boero (Eds.), *The second handbook of research on the psychology of mathematics education* (pp. 151-185). Sense Publishers.

Schmidt, W., Wang, H., & Mcknight, C. (2005). Curriculum coherence: An examination of US mathematics and science content standards from an international perspective. *Journal of Curriculum Studies*, *37*(5), 525–559.

Stein, M. K. (2000). *Implementing standards-based mathematics instruction: A casebook for professional development*. Teachers College Press.

Watt, H. (2005). Attitudes to the use of alternative assessment methods in mathematics: A study with secondary mathematics teachers in Sydney, Australia. *Educational Studies in Mathematics*, *58*(1), 21–44.

Wellberg, S. (2023). Teacher-made tests: Why they matter and a framework for analysing mathematics exams. *Assessment in Education: Principles, Policy & Practice*, *30*(1), 53-75.

Wiggins, G. (1992). Creating tests worth taking. *Educational Leadership*, *49*(8), 26–33.

Zhang, Q., Zhang, X., & Liu, J. (2021). A holistic review of authentic assessment in mathematics education. In T. Barkatsas & P. McLaughlin (Eds.), *Authentic assessment and evaluation approaches and practices in a digital era: A kaleidoscope of perspectives* (pp. 95-115). Brill Publisher.

HOW DOES MATHEMATICAL CREATIVITY IN ALGEBRA CHANGE ACROSS SECONDARY UNDER STUDENT- CENTERED AND TEACHER-CENTERED PEDAGOGY?

Ying Zhang

University of Cambridge

This study explores the developmental trajectory of mathematical creativity within secondary students, and whether this trajectory differs between pedagogy. A comparative case study of two Chinese secondary schools (Grades 7-9) was conducted, which in our prior research we found differ significantly in their delivered pedagogy: one is more student-centered pedagogy and the other more teacher-centered pedagogy. Using cross-sectional data, this study conducted within- and between-school comparisons at the beginning of Grade 8 (N=182) and at the end of Grade 9 (N=162). Notable findings included significant differences between the creativity of Grade 8 and Grade 9 students, with the latter group demonstrating creativity that was twice as high. This trend applies for both schools, regardless of the pedagogy students received.

INTRODUCTION

Creativity plays a crucial role in the full cycle of advanced mathematical thinking. Mere mastery of mathematical material is not a sufficient criterion for mathematical giftedness, but needs to be extended to an “independent creative mastery of mathematics under the conditions of school instruction” (Krutetskii, 1976, p.68). Notably, creativity is not a static entity but develops as people mature, and students’ mathematical creativity can vary and develop across grade levels (e.g., Cheung et al., 2004). Understanding these changes would thus yield significant insights for researchers seeking effective ways to foster creativity. Despite decades of studies, there is still a lack of clarity regarding the developmental trend of creativity during elementary and secondary education. For example, Odelya (2023) reported that as the age of students rises, they are less prone to looking for creative solutions and more likely to be “held hostage” by their habitual use of algebra. However, Cheung et al. (2004) reported secondary students’ creativity rose from Grade 1 to Grade 9.

On the other hand, regarding ways to foster mathematical creativity, some researchers suggested that, in contrast with teacher-centered pedagogy (TCP), student-centered pedagogy (SCP) has such a potential (e.g., Torrance, 1966). Yet no significant differences were found between the SCP and TCP Grade 9 students when Zhang (2023) explored such relationship for a preliminary attempt, which might be attributed to the limited tasks or the time-length students have received the pedagogies. Also, it is unknown whether the developmental trend of creativity differs between TCP and SCP. Hence, the relationship among creativity, pedagogy, and grade levels needs to be further investigated. Considering the research gaps mentioned above, this study moves

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. 4, pp. 233–240). PME.

a step towards addressing these needs for research by focusing on the following research question: Are there significant differences between the mathematical creativity of the younger and older junior secondary students depending on whether they experienced SCP and TCP?

THEORETICAL FRAMEWORK

Teacher-centered pedagogy (TCP) is often described as being based upon a model of an active teacher and a passive student; in contrast, student-centered pedagogy (SCP) is based upon the idea of an active student. In mathematics, SCP has been used to describe a learning environment where (a) student mathematical thinking is made public, (b) students actively engage with each other's mathematical thinking, and (c) student mathematical sense-making, conjecturing, and justifying drive instruction (Thanheiser & Melhuish, 2023).

This study is concerned with the "little-c" creativity, the relative creativity of non-experts and students as they generate novel mathematical ideas that are new to the students' previous experiences or the performance of other students with similar educational history (e.g., Leikin, 2009). Two views of the "little-c" have been provided in the literature: The first one considers that creativity includes not only convergent thinking, the ability to generate a single correct solution to a problem, but also divergent thinking, the thought process used to generate multiple possible solutions to a problem (Guilford, 1967). The second view considers creativity based on fluency, flexibility, and originality (Torrance, 1966): *fluency* refers to the continuity of ideas and flow of associations; *flexibility* refers to the variety of approaches to a problem; and *originality* is characterized by a unique way of thinking and respective unique products of mental activity. The notion of creativity in our study aligns with both of these views whose integration can be found in multiple-solution tasks (MSTs), which examine both divergent and convergent thinking as suggested by Guilford (1995) and as reflected in problem solving processes and outcomes (Leikin, 2009), and are measured via the three components suggested by Torrance (1966).

METHODOLOGY

Comparative case study

Dulangkou Secondary School and School Y (pseudonym) were selected as cases of schools implementing primarily SCP and TCP, respectively. The two junior secondary schools were selected due to having comparable features but different pedagogical approaches. Regarding similarities, firstly, both schools are in rural towns under the same county of the same city, so they follow the same educational policies and have similar economic conditions. Secondly, both schools randomly divide students into classrooms rather than based on achievement. Thirdly, they are the only schools in their respective towns, both of which require their school to recruit students only from within the district; thus, the two schools have similar student-intake processes.

Regarding the differences, School Y is one of the best-performing schools among all 14 towns that purportedly employs TCP, while Dulangkou is the most popular school among all towns due to its reformed SCP (Zhang & Stylianides, 2023), which exemplified the result of Chinese compulsory education reform. This study used the RTOP observation protocol (Piburn et al., 2000), which was developed to evaluate the extent to which a classroom adopts reform-based pedagogy, to investigate and ultimately confirm that Dulangkou uses more SCP and School Y more TCP. Four Dulangkou (11 lessons) and five School Y mathematics teachers (13 lessons) were observed in 2022. The RTOP scale ranges from 0 to 100, where lower scores reflect TCP and higher scores reflect SCP. A Mann-Whitney U test showed School Y scored significantly lower than Dulangkou on the RTOP scale (Zhang & Stylianides, 2023). In light of these findings, this paper uses "SCP school" and "TCP school" to represent Dulangkou and School Y, respectively.

Grade 7 is the first year, and Grade 9 is the last year, of Chinese junior secondary education. Within each school, students at the beginning of Grade 8 and the end of Grade 9 were selected as cases of grade levels. Consequently, a nearly two-year disparity in learning duration exists among the participants from these two grade levels. In total, two classrooms of Grade 8 (97 SCP, 85 TCP) and Grade 9 (83 SCP, 79 TCP), respectively, from each school were randomly selected to participate.

Methods

This study used multiple-solution tasks (MSTs) to indicate mathematical creativity, which are open-ended tasks explicitly required students to solve a mathematical problem in different ways (Leikin, 2009). The task, described in Table 1, was chosen for the following three reasons. Firstly, it is a Grade 7 task appearing often in past Chinese examinations, assuring their content validity in terms of problem solving. Secondly, this task is at a relatively easy level to prevent students' divergent thinking to be submerged by problem-solving skills. Otherwise, MSTs would be degraded into assessing problem-solving competencies rather than creativity (Zhang, 2023). Thirdly, the knowledge covered at the task is routine and is emphasized by the high school entrance examination, thus students would regularly encounter it through their secondary since Grade 7. In light of this, the study satisfies the precondition outlined by creativity, that participants have similar background with respect to the task, which is not biased towards one school over another or more familiar to a specific grade level.

Table 1. Multiple-solution task used in this study

Solve the following problem in as many ways as possible: Guihua city dispatched a total of 15 citizens, both male and female, to purchase cement for construction and carry them back. It is known that every male citizen carried two bags of cement, while two female citizens carried one bag. In total, they have purchased and carried 15 bags of cement back. How many male and female citizens were dispatched for this procurement, respectively?

All tests were administrated in person by the author, with students' desks arranged separately during their regular class time. To ensure that students treat the assessment seriously, they were informed that the task will be similar to those found on the coming High School Entrance Examination, and thus they should view this assessment as an opportunity to test their potential.

Solutions generated by students were first marked based on appropriateness. The notion of appropriateness allows evaluating reasonable ways of solving a problem that potentially led to the correct solution outcome regardless of the minor mistakes made by a solver (Leikin, 2009). The data was then analyzed based on the creativity rubric adapted from Leikin (2009), the detailed of which has been described in Zhang (2023). Finally, students' responses were categorized into the Creative Thinking Level (CTL), adapted from Siswono (2011), in alignment with Leikin's (2009) rubrics as described in Table 2. The following ordinal logistic regression model was employed to analyze the CTL results and to obtain creativity odds ratio: logit $(P(Y \leq k|S))=\log_e(\frac{P(Y\leq k|S)}{1-P(Y\leq k|S)})$.

Table 2. Creative Thinking Level adapted from Siswono (2011)

Level	**Characteristic of Creative Thinking Level**
Level 0 (Not Creative)	Students were not able to show any components of creativity ($Cr = 0$)
Level 1 (Almost Not Creative)	Students were able to show fluency with low originality and flexibility in solving problem ($Flu_i \geq 1$, $Flx_{i\neq 1} < 10$ and $Ori_i \leq 1$)
Level 2	Students were able to show flexibility with low fluency ($0 < Flu_i < 2$, $Flx_{i\neq 1} \geq 5$, $Ori_i \leq 1$)
Level 3 (Quite Creative)	Students were fluent and flexible ($Flu_i > 1$, $Flx_{i\neq 1}=10$ and $Ori_i < 10$)
Level 4	Students were able to show originality in solving problem with low fluency and flexibility ($0 < Flu_i \leq 1$, $Flx_{i\neq 1} < 10$ and $Ori_i = 10$)
Level 5 (Creative)	Students were fluent and they demonstrate originality ($Flu_i > 1$, $Flx_{i\neq 1} < 10$ and $Ori_i = 10$)
Level 6 (Very Creative)	Students satisfied all components of creativity ($Flu_i > 1$, $Flx_{i\neq 1}=10$ and $Ori_i=10$)

RESULTS

Fluency

Table 3 categorizes the fluency score, Flu_i, generated by students. For example, 16.5% SCP Grade 8 scored between 0 and 1 (inclusive). Interestingly, the mean fluency scores at the two schools are identical for both grade levels (0.68 for Grade 8, 1.42 for Grade 9). The Grade 9 students at both schools showed higher percentages of students, not only in solving the task but also in generating two, three, and four solutions. Via both Welch's T test and Mann-Whitney U test, Grade 9 participants significantly outperformed ($p<0.05$) those in Grade 8 in terms of fluency.

Table 3. Number of solutions corresponding to percentage of students

$\mathrm{Flu_i}$	0	(0, 1]	(1,2]	(2,3]	(3,4]	(4,5]	Mean $\mathrm{Flu_i}$
Grade 8							
SCP	59.8%	16.5%	18.6%	4.1%	1.0%	0%	0.68
TCP	50.6%	31.8%	15.3%	1.2%	1.2%	0%	0.68
Grade 9							
SCP	32.5%	19.3%	28.9%	9.6%	8.4%	9.6%	1.42
TCP	34.2%	19.0%	29.1%	7.6%	6.3%	10.1%	1.42

Flexibility

Table 4 shows the percentage of students generating the corresponding strategy within their school participants. For example, 1.0% SCP participants and 1.2% TCP participants used strategy D. The SCP Grade 8 covered two less categories than the ones of the TCP Grade 8, while the SCP Grade 9 covered one more category than the ones of the TCP. Both TCP grade levels had higher percentages of students using Strategy C and D, whereas both SCP grade levels had more percentages using Strategy F. Notably, the Grade 8 students at both schools had Strategy A as the greatest number of responses, while the results switched to Strategy B for Grade 9 at both schools.

Table 4. Distribution of categories of solutions (Grade 8 left, Grade 9 right)

Grade 8	SCP	TCP	$\mathrm{Ori_i}$
Strategy A	25.3%	27%	0.1
Strategy B	6.2%	8.3%	1
Strategy C	0	1.2%	10
Strategy D	1.0%	1.2%	10
Strategy E	9.2%	9.0%	1
Strategy F	21.6%	16.7%	1
Strategy G	0	1.2%	10

Grade 9	SCP	TCP	$\mathrm{Ori_i}$
Strategy A	39.8%	43.0%	1
Strategy B	56.6%	49.4%	1
Strategy C	2.4%	8.9%	10
Strategy D	6.0%	14.0%	10
Strategy E	7.2%	8.9%	10
Strategy F	16.9%	6.3%	10
Strategy G	1.2%	0	10

Originality

The corresponding originality values, $\mathrm{Ori_i}$, for every strategy are presented in Table 4. Notably, Strategy A, E, and F received higher originality within Grade 9 than within Grade 8. Table 5 lists the percentage of students generating original solutions and the mean originality achieved by each group. For example, 17.5% of the Grade 8 SCP participants had at least one of their solutions scoring 10.0 originality, the highest one can achieve, and the mean originality achieved by Grade 8 SCP students was 0.83. Notably, the Grade 9 students from both schools achieved higher originality scores than those of Grade 8; both grade levels of the TCP group achieved higher mean originality and had more participants generating original solutions than the ones of the SCP group.

Table 5. Mean originality and percentage of students generating original solutions

	Grade 8 SCP	Grade 8 TCP	Grade 9 SCP	Grade 9 TCP
Total Ori_i	17.5%	18.8%	26.5%	30.4%
Mean Ori_i	0.83	1.00	4.40	5.09

Creativity

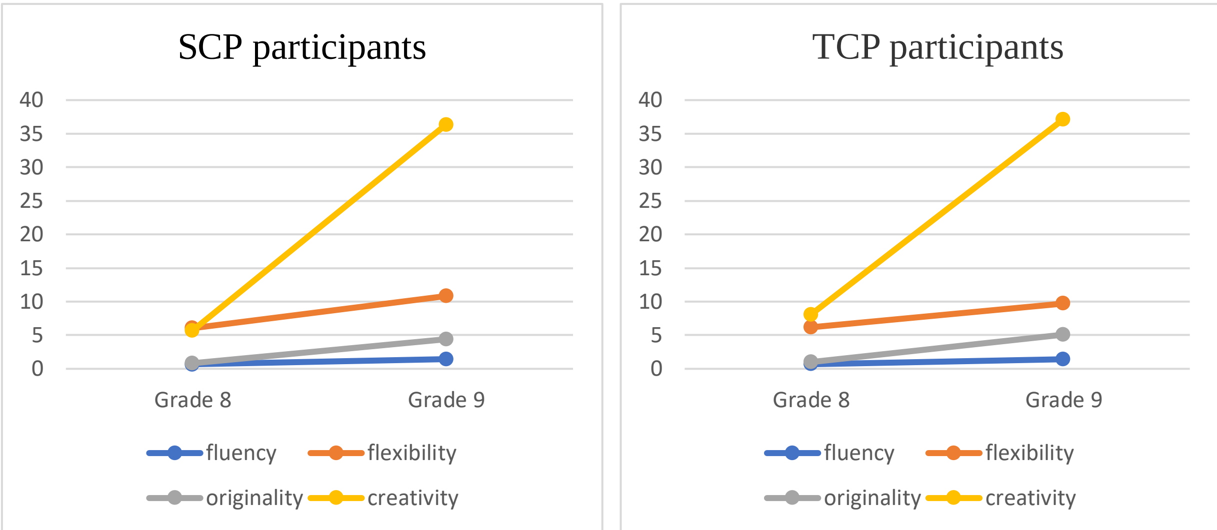


Figure 1. Average scores for each cognitive construct across grades

A Mann-Whitney U test was used to evaluate the hypothesis that Grade 9 from each school would score higher, on the average, than Grade 8 on total creativity scores. The results were in the expected direction (Figure 1) and significant, $p<0.05$. Yet no significant differences were detected between the creativity of the two schools for any grade levels. Remarkably, although the SCP Grade 9 participants had higher mean fluency and flexibility, its mean originality and creativity scored lower than the respective groups of the TCP. This suggests that originality can be independent from fluency and flexibility, and originality plays a dominating role in creativity.

The CTL results, via logistic regression, suggest that the odds of Grade 9 students obtaining a higher CTL on the algebraic task is 2.72 as it is for Grade 8 students, with 2.76 for the SCP school and 2.67 for the TCP school. The odds of the SCP Grade 8 and Grade 9 students, respectively, obtaining a higher CTL on the given task is 0.90 times and 0.98 times as large as it is for the TCP Grade 8 and Grade 9 students. The differences between grade levels are larger than it between schools, and the creativity differences between the two schools were narrowed down marginally in Grade 9.

DISCUSSION

All four constructs including creativity increased significantly from Grade 8 to 9 across both schools (Figure 1). This implicates that, at the stage of junior secondary, students' mathematical creativity increases along with their grade levels, regardless of the pedagogies received. This trend is consistent with Cheung et al. (2004), who reported an increasing creativity from Grade 1 to 9 via Wallach-Kogan Creativity Tests, lending credence to the cognitive notion that older children are supposed to perform at a higher level of creativity as their social experiences and educational training become broader (Cheung et al., 2004). However, this finding contradicts Odelya and Miriam (2023), who found that students are less prone to looking for creative solutions as the age rises. Their results may be highly influenced by the non-routine nature of the task, within which students from different grade levels hardly have a consistent educational background and thus the task selected may be more favorable for the younger students. This inconsistency also suggests that students' creativity in routine and non-routine tasks could have a different developmental trajectory.

Both grade levels at the TCP school possessed higher originality and creativity scores than the SCP group, indicating that students from teacher-centered environment can possess same level of or even higher creativity in algebra. This might seem counterintuitive but aligns with the creativity tests results conducted by PISA in 2022 on 15-year-old students (OECD, 2023), where the top six positions were dominated by East Asian regions, which was often presumably to be more teacher-centered. Specifically, three of the top six derive from Chinese regions (OECD, 2023).

Align with Zhang (2023), no significant differences were found between the creativity of the two schools on the given algebraic task, suggesting that the role pedagogy plays in algebraic creativity might be smaller than it can be detected, and the progress linked to grade levels or long-term learning duration may surpass that of pedagogy in terms of its impact on creativity development. This implies that creativity may require sufficient familiarity with a content and even higher problem-solving competencies, and mathematical creativity occurs after the algorithmic fixation in problem solving, where fixation is shown in the repeated use of an initially successful algorithm or procedure (Krutetskii, 1976).

This study also found that the most conventional strategy, along with the originality values for any strategy, can both vary across grade levels. This could be attributed to the increasing problem-solving skills of Grade 9 students, more of whom were able to generate the most conventional strategy, not only increasing the fluency scores but also making the other non-trivial solutions rarer in terms of the percentage. Given this, the relationship between mathematical problem solving and creativity should be further investigated by future research.

Limitations

The limited number of the MSTs used in the study may not fully capture the entire range of students' mathematical creativity in algebra. Also, although this study endeavored to have pedagogy be the main variable influencing creativity, no causal claim can be made between pedagogy and creativity owing to the limitation of comparative case studies: the quandary of "many variables, small-N".

References

Cheung, P. C., Lau, S., Chan, D. W., & Wu, W. Y. H. (2004). Creative potential of school children in Hong Kong: Norms of the Wallach–Kogan creativity test and their implications. *Creativity Research Journal*, 16, 69–78.

Guilford, J. P. (1967). Creativity: Yesterday, today and tomorrow. *The Journal of Creative Behavior*, *1*(1), 3-14.

Krutetskii, V. A. (1976). In J. Kilpatrick, & I. Wirszup (Eds.). *The psychology of mathematical abilities in school children*. Chicago: University of Chicago Press.

Leikin, R. (2009). Exploring mathematical creativity using multiple solution tasks. In *Creativity in mathematics and the education of gifted students* (pp. 129-145). Brill.

Odelya, U., & Miriam, A. (2023). Non-routine problem solving and creativity among talented math students from a multi-age perspective. 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, p.343). Haifa, Israel.

OECD (2023). PISA 2022: Insights and Interpretation. COVER FINAL.indd (oecd.org)

Piburn, M., Sawada, D., Turley, J., Falconer, K., Benford, R., Bloom, I., & Judson, E. (2000). *Reformed teaching observation protocol (RTOP) reference manual*. Tempe, Arizona: Arizona Collaborative for Excellence in the Preparation of Teachers.

Siswono, T. Y. E. (2011). Level of student's creative thinking in classroom mathematics. *Educational Research and Reviews*, *6*(7), 548-553.

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*, 101032.

Torrance, E. P. (1966). Torrance tests of creative thinking. *Educational and Psychological Measurement*.

Zhang, Y. (2023). Exploring the role of pedagogy in mathematical creativity via multiple-solution tasks: A comparative study of two schools in China. In M. Ayalon, B. Koichu, R. Leikin, L. Rubel., & M. Tabach (Eds.), *Proceedings of the 46th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp.379-386). Haifa, Israel.

Zhang, Y., & Stylianides, A. J. (2023). A comparative case study of the mathematics pedagogy in two Chinese schools: How "student-centered" is a proclaimed reformed pedagogy? In, *Proceedings of the Thirteenth Congress of European Research in Mathematics Education*. Budapest, Hungary.

CHINESE STUDENTS' MATHEMATICAL WELLBEING THREE YEARS ON: A RE-ASSESSMENT IN GRADE 6

Juan Zhong[1], Veysel Akçakın[2], and Wee Tiong Seah[3]

[1]Jinsha Elementary School, China; [2]Uşak University, Türkiye;
[3]The University of Melbourne, Australia

The mathematical wellbeing (MWB) of 76 students in a suburban elementary school in Chengdu, China were assessed twice, once in 2020 when they were part of a bigger Grade 3 participant group, and again in 2023 when they were in Grade 6. The same questionnaire was used, with its presentation adjusted to match students' ages. Variable/facet parameters were determined using Many Facet Rasch Measurement, and the Rasch-Welch t-test was employed to compare differences between Grades 3 and 6. Analysis found that the fulfilment of the same values contributed to students' MWB at both grade levels. However, at Grade 6, MWB was associated with more experiencing of the valuing of accomplishment and perseverance, less experiencing of engagement and bliss, and similar levels of relationship and meaningfulness.

INTRODUCTION

Given the enabling effect of general wellbeing on human flourishing (Chaves, 2021), the fostering of mathematical wellbeing (MWB) amongst students can promote effective mathematics learning while reducing the likelihood of disengagement and mathematics anxiety. While MWB (and other affective traits) might be cultivated in early and elementary school years, we are concerned that it might be eroded as students progress up the grade levels. Especially since MWB is an expression of the extent to which relevant values are fulfilled, how might such values fulfilment be affected by mathematics topics and/or pedagogies in upper elementary or high school curricula, which would in turn impact on MWB?

This paper reports on a study conducted with a group of Grade 6 students in the Chinese city of Chengdu, whose MWB had been assessed in 2020 in a previous study when they were in Grade 3, and which was assessed again in their final year of elementary schooling (i.e., Grade 6) in 2023. Thus, this study design incorporates the advantage of surveying from the same students a few years apart, rather than making inferences from surveying students of different grade levels at any one time period.

MATHEMATICAL WELLBEING (MWB)

We regard MWB as being "the fulfilment of core values … within the mathematics learning experience, accompanied by positive feelings (e.g., enjoyment, pride) and functioning (e.g., accomplishment, engagement) in mathematics" (Hill & Seah, 2023, p.386). Developing and maintaining positive MWB amongst students is important not just because mathematics is one of a few subjects that is studied by all students

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. 4, pp. 241–248). PME.

globally, but also because so many students experience disengagement in – or negative attitude to – mathematics lessons, and/or mathematics anxiety. Intervention approaches can be costly and yet success is not guaranteed. On the other hand, if we proactively develop and maintain students' MWB, more students around the world can get to learn mathematics with positive affect, as well as effectively, to help them navigate the complexities and uncertainties of our current world.

Data collected and analysed in Australia, China and New Zealand had validated a set of seven ultimate values the fulfilment of which is needed to achieve MWB (Hill et al., 2022). These values are *accomplishment, cognitions, engagement, meaning, perseverance, positive emotions,* and *relationship*. While these ultimate values might be the same across cultures, the instrumental values serving them have been found to be different (Hill & Seah, 2023).

THE PREVIOUS STUDY

The 'previous study' mentioned above refers to a similar study (Pan et al., 2022) conducted in 2020 when the same student participants were in Grade 3 in the same school. In fact, they were part of a larger group of 258 Grade 3 students in six classes in the Chengdu suburban school, taught by three mathematics teachers. The teachers had nominated 21 classroom learning moments (e.g., 'when you are given an interesting mathematics learning task', 'when your mathematics teacher praises you) to which students indicated the extent to which they valued each and were able to live it. There was also an 'other' option for students to indicate classroom learning moments associated with positive MWB. Engaging in these learning moments enabled the students to fulfil and live some or all of 15 instrumental values (Figure 2). The instrumental values together would serve the realisation of a smaller set of 6 terminal or ultimate values (Figure 1). For example, the learning moment 'when you are given an interesting mathematics learning task' was considered to help students fulfil their instrumental valuing of *interestingness*, which was in turn considered to serve the ultimate valuing of *engagement*.

Three findings were of particular importance in this previous study. Firstly, the students' MWB corresponded to the fulfilment of a set of seven ultimate values which are similar to the set that Hill et al. (2021) observed in Australia, namely, *relationship, engagement, bliss, accomplishment, perseverance, meaningfulness,* and *learning.* Secondly, four of these - *engagement, relationship, bliss* and *accomplishment* – were especially emphasised by the students for positive MWB. Thirdly, teachers' facilitation of these values which fostered positive MWB was generally consistent across different teachers and different classes.

Given that student affect often becomes less positive as they progress through grade levels (e.g., Thomson et al., 2020), this current study is interested to find out what the MWB of some of these 258 students were like in their final year of elementary schooling. In particular, the Research Questions guiding the conduct of this study are:

RQ1: What are the ultimate values that need to be fulfilled in order for Grade 6 students in Chengdu to experience mathematical wellbeing? How do these compare with the ultimate values associated with these students when they were in Grade 3?

RQ2: For each of the ultimate values associated with Grade 6 students' mathematical wellbeing, how similar or different are the corresponding instrumental values compared to the time when the students were in Grade 3?

METHODOLOGY

Participants

The student participants in this study were 76 Grade 6 students in two classes in a Chengdu suburban elementary school. They (and their mathematics teacher) were also part of the 258 participants who provided data in 2020, when they were in Grade 3. They have had the same mathematics teacher throughout the six years of elementary education in the same school, but the different mathematics topics and the different pedagogies associated should affect individuals' mathematics learning experience. For example, as mathematics topics become more abstract in the upper elementary school year levels, and as different teaching approaches need to be introduced, how might these affect the extent to which students were able to engage in classroom learning moments that reflect the fulfilment of relevant values? How might these affect MWB?

The Questionnaire Method

Just like when they were in Grade 3 three years before the current study, the students' MWB was assessed through the questionnaire survey method. Compared to alternative methods such as interviews and journals, the questionnaire approach would have facilitated efficient collection of data from a large group of participants at the same time. In both times, the students completed the questionnaires during mathematics lesson time, with the same mathematics teacher administering the exercise.

The questionnaire (in Chinese) is accessible at https://www.wjx.cn/vm/POhZjXH.aspx. While the items are the same as the questionnaire which the student participants completed three years prior (see Hill & Seah, 2023), there were necessarily some changes in the way it was presented, considering that the students had become older and more matured. Firstly, students had indicated in the earlier questionnaire if 21 given learning moments were associated with positive MWB through a colouring exercise. The argument then was that the activity would help maintain the young students' attention span. In the current questionnaire, students only needed to click on bullet points adjacent to the 21 learning moments statements to indicate that they were still associated with their experiencing of positive MWB. Like the Grade 3 questionnaire, there was an additional 'other' option too. (For a list of the 21 learning moments, refer to the English translated version of the questionnaire, accessible at: https://melbourneuni.au1.qualtrics.com/jfe/form/SV_doQ5pV3ruEwyZTw) Secondly, while the questionnaire was administered in hardcopy version in the earlier exercise in 2020, it was presented to students as an online survey in the current exercise in 2023.

The questionnaire responses were exported in the form of a Microsoft Excel spreadsheet. The content (i.e., raw data) were cleaned and organized in preparation for Many Facet Rasch Measurement [MFRM] analysis. The MFRM is a measurement model in the item response theory that extends the Rasch model (Toffoli et al., 2016). Thus, the codes were written as guided by the FACET software to facilitate our investigation of the interaction between grade level and instrumental / ultimate values.

With the Research Questions listed above in mind, we focussed on Item 6 of the questionnaire. Each student was scored according to whether each of the 21 learning moments contributed to their MWB, based on their indications in the Grade 6 questionnaire. The same question in the Grade 3 questionnaire, however, gave the students three choices of which to colour-in one: contributes a lot, contributes, and does not contribute. In our analysis, responses to either one of the first two choices were counted together. In other words, all student responses were recorded as either 1 or 0, thus implying that the data were dichotomous.

In contrast to classical test theory, MFRM allows for the independent and objective estimation of facet parameters without being influenced by item, rater, test, and group characteristics (Toffoli et al., 2016). In this study, individuals (students), grade level, instrumental values, ultimate values, and learning moments were determined as facets. MFRM enables the separate estimation of each facet and allows for relevant comparison by calibrating and standardizing the facets within a logit scale (Prieto et al., 2014), where scores generally fall between -3 and +3.

Item response theory is relevant in our study as it helps us to examine the relationship between the latent MWB and the observable learning moments which made up the item responses. With the dichotomous data coded, the Rasch-Welch t-test was performed to compare the difference between ultimate values experienced in Grades 3 and 6, because this test is more effective in controlling Type 1 error rates when the assumption of equal variance is not fulfilled (such as in this study), while maintaining a strong level of reliability compared to Student’s t-test if the assumptions are met (Delacre et al., 2017).

An informal conversation was also set up with the classroom teacher to share with her what the analysed data looked like, to stimulate her thoughts and opinions in response.

RESULTS

Point-measure (point-biserial) correlation values of the items for the 21 learning moments vary between .43 and .74. Infit MNSQ values range from .82 to 1.23, and outfit MNSQ values range from .59 to 1.28 except for one item that is .44. These values being between .5 and 1.5 are productive in terms of measurement. Values lower than 0.5 are not as productive for measurement, but they do not cause degradation. (Linacre, 2002). These results show that our data fit the Rasch model.

The interaction of grade level and ultimate values is shown in Figure 1, while the interaction of grade level and instrumental values is shown in Figure 2.

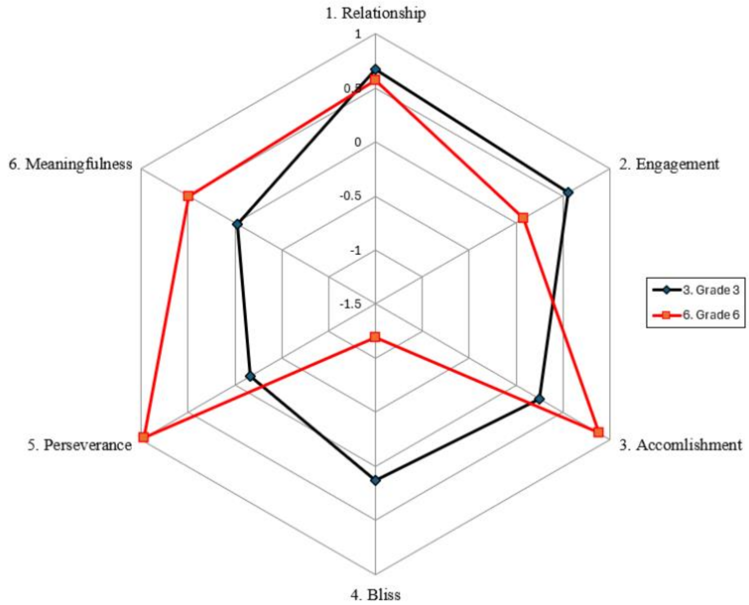


Figure 1: Interaction of Grade Level and Ultimate Values.

Rasch-Welch (logistic regression) t-test results show that there are statistically significant differences between the Grade3 and Grade 6 students in favour of the Grade 3 students for the ultimate values of engagement ($t_{(1115)}$=3.02, p<.05) and *bliss* ($t_{(273)}$=4.41, p<.05), and in favour of the Grade 6 students for the ultimate values of *accomplishment* ($t_{(882)}$=-3.62, p<.05) and *perseverance* ($t_{(347)}$=-4.20, p<.05).

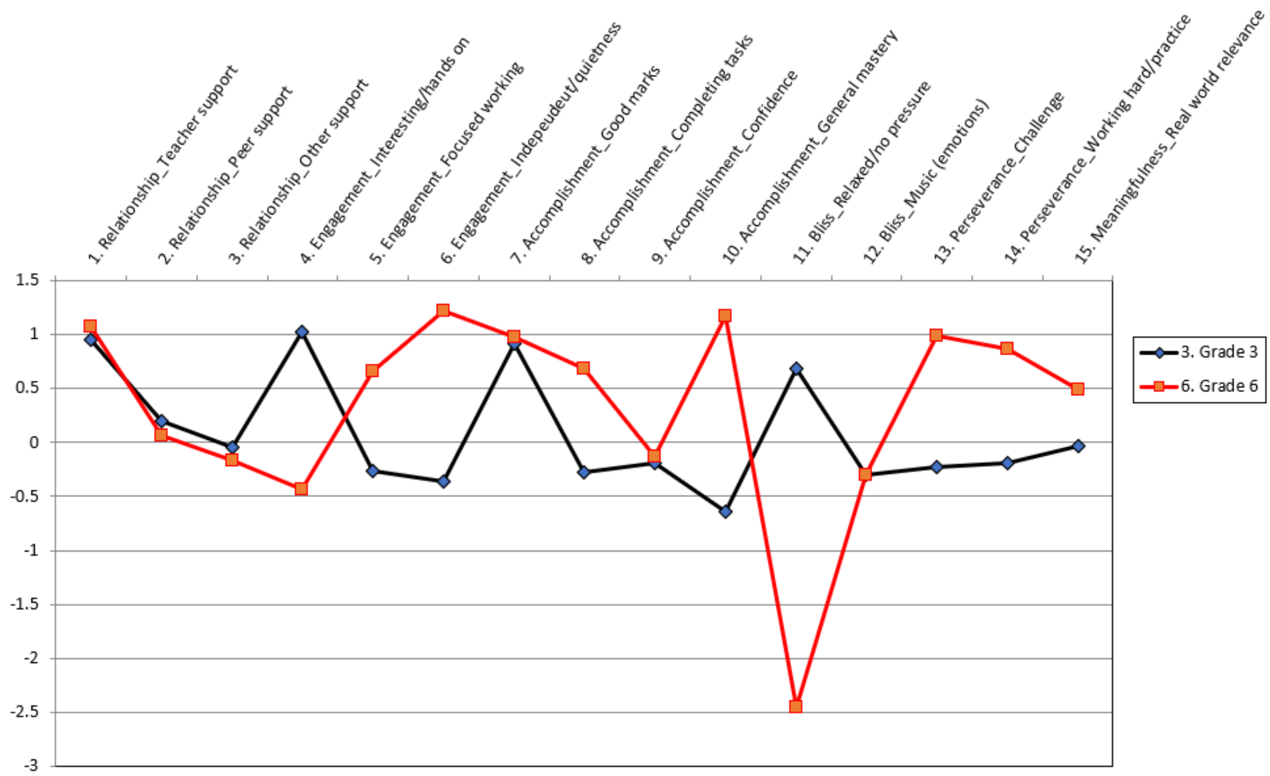


Figure 2: Interaction of Grade Level and Instrumental Values.

Rasch-Welch (logistic regression) t-test results show that there are statistically significant differences between the Grade 3 and Grade 6 students in favour of the Grade 3 students for the instrumental values associated with ‘interesting/hands-on’ ($t_{(758)}=6.97$, $p<.05$) and ‘relaxed/no pressure’ ($t_{(110)}=5.66$, $p<.05$); and in favour of the Grade 6 students for the instrumental values associated with ‘focused working’ ($t_{(172)}=-2.52$, $p<.05$), ‘independent/quietness’ ($t_{(166)}=-4.09$, $p<.05$), ‘completing tasks’ ($t_{(174)}=-2.58$, $p<.05$), ‘general mastery’($t_{(162)}=-4.78$, $p<.05$), ‘challenge’($t_{(171)}=-3.16$, $p<.05$), and ‘working hard/practice’ ($t_{(173)}=-2.78$, $p<.05$).

DISCUSSION

76 Grade 6 students in a Chengdu suburban elementary school were asked to identify, from a teacher-nominated set of learning moments, those which accompany their experiencing of MWB. A similar assessment was carried out with this group of students three years prior when they were in Grade 3. In responding to Research Question 1, it was found that at this upper elementary level, MWB was associated with the fulfilment of six of the seven ultimate values identified earlier, that is, without *learning*. This was to be expected, since none of the 21 teacher-nominated learning moments reflected the valuing of *learning*, and the reason why it was an ultimate value associated with MWB three years prior was that two students (of the 258) then had identified them in the open-ended ‘other’ item. This is not to suggest, however, that students’ MWB did not involve experiencing of *learning*: conversations with the mathematics teacher suggest that the students were not short of opportunities to experience the valuing of *learning*. In other words, even though learning moments reflecting *learning* might have been too obvious for the classroom teachers to have listed them in the questionnaire, this current study lends further support for the same set of seven ultimate values governing MWB as was identified in Hill et al. (2022).

Specifically, over the three years from Grade 3 to Grade 6, two each of the six ultimate values were associated more with MWB and experienced more by students (*accomplishment, perseverance*); less associated and experienced (*engagement, bliss*); and similarly associated and experienced (*relationship, meaningfulness*).

Research Question 2 aimed to understand which instrumental values experienced changes in fulfilment that led to changes in the fulfilment of the associated ultimate values. The statistically significant drop in students’ experiencing of *bliss* in Grade 6, for example, could be the result of a drop in the fulfilment of being *relaxed* (one of two instrumental values assessed), whereas the fulfilment of (listening to) *music* (the other instrumental value assessed) remained the same over the three years. Similarly, the increase in students’ fulfilment of *accomplishment* and *perseverance* was due to changes in two instrument values each: *completing tasks* and *general mastery* for the former, and *working hard* and *challenge* for the latter. Notably, the drop in expression of students’ *engagement* seemed to be caused by all three instrumental values assessed, namely, *interestingness*, *focussed work*, and *independence*.

Despite the changing nature of mathematics topics at upper elementary levels, despite the demands and needs of adolescence, the data suggest that the students’ experiencing of their valuing of *relationship* has not been affected. Perhaps this is because students in China have the same teachers and peers throughout their elementary school years. Students' experiencing of *meaningfulness* has also remained stable.

It is not surprising that even as *bliss* continued to be a value underlying MWB, Grade 6 students were experiencing less of it. Not only have mathematics topics become more difficult (and abstract), parental pressure on results, teachers having less opportunities for positive reinforcements, and more complex question types all contributed. The changing nature of classroom activities away from fun ones such as origami (in Grade 3) probably also explained the less fulfilment of *bliss* and *engagement* at Grade 6.

The classroom teacher was aware that her students' opportunities to experience the valuing of *engagement* were being threatened. In response, she introduced group-based mathematics projects to her students annually, recognising that these would stimulate students' interest in hands-on tasks, promote focussed working, and provide students with the independence they enjoyed in completing the respective projects. These three aspects are in fact the instrumental values (see Figure 2) promoting student engagement in their mathematics learning. Yet, the projects probably did not exert sufficient influences to the students' engagement. Another point to note is that at the time of collecting the Grade 6 data in 2023, the year's project had not been announced yet.

The classroom teacher had been surprised that her students were fulfilling *accomplishment* and *perseverance* more, when she was expecting these to slide in Grade 6. According to her, this concern had probably led her to over-compensate, by consciously building into her lessons more opportunities for students to exercise perseverance, and to feel accomplished. This suggests that intentional teacher actions in their professional practice can effectively affect values fulfilment, and thus, MWB.

The data suggest that as students progress through the elementary school years in China, the development / maintenance of their MWB does not require the fulfilment of different values. However, the changing nature of the curriculum and the changing preferences of growing children have meant that opportunities for relevant instrumental values – and thus, the learning moments in class – to be experienced by the students are different across grade levels. Teacher awareness of these are important, for as the mathematics teacher in this study showed, teachers can use this knowledge to orchestrate student experiencing of targeted learning tasks to facilitate the fulfilment of particular values. Furthermore, the learning moments are commonly found in mathematics classrooms, implying that teachers need not introduce intervention activities into their lessons, disrupting established lesson structures.

CONCLUSION

This paper reports on the second assessment of students' MWB for a longitudinal study in an elementary school in Chengdu, China. Three years on after the first assessment in Grade 3, the Grade 6 students' MWB were supported by six ultimate values which were also documented three years prior, namely, *accomplishment*, *perseverance*, *meaningfulness*, *relationship*, *bliss*, and *engagement*. A seventh value, *learning*, was neither surveyed nor identified by the Grade 6 students, although we were not surprised when the classroom teacher believed that students' experiencing of it would also contribute towards their MWB. Amongst the six identified values, the students reported experiencing more of the first two ultimate values three years on, equivalent experience with the middle two, and less experiencing of the last two. The instrumental values underlying these changes were identified, with students experiencing less of all the three instrumental values feeding into *engagement* in particular.

References

Chaves, C. (2021). Wellbeing and flourishing. In: M.L. Kern, & M.L. Wehmeyer (Eds.), *The Palgrave Handbook of Positive Education*. Palgrave Macmillan. https://doi.org/10.1007/978-3-030-64537-3_11

Delacre, M., Lakens, D., & Leys, C. (2017). Why psychologists should by default use Welch's t-test instead of Student's t-test. *International Review of Social Psychology, 30*(1), 92-101. https://doi.org/10.5334/irsp.82

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 well-being and values. *ECNU Review of Education*, *4*(2), Article 2. https://doi.org/10.1177/2096531120928084

Hill, J. L., Kern, M.L., Seah, W.T., & van Driel, J. (2022). Developing a model of mathematical wellbeing through a thematic analysis of the literature. In C. Fernández, S. Llinares, A. Gutiérrez, & N. Planas (Ed.), *Proceedings of the 45th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 379–386). PME.

Hill, J. L., & Seah, W. T. (2023). Student values and wellbeing in mathematics education: Perspectives of Chinese primary students. *ZDM – Mathematics Education*, *55*(2), 385–398. https://doi.org/10.1007/s11858-022-01418-7

Linacre, J. M. (2002). What do infit and outfit, mean-square and standardized mean? *Rasch Measurement Transactions*, *16*, 871-882.

Pan, Y., Zhong, J., Hill, J. L., & Seah, W. T. (2022). Values which facilitate mathematical wellbeing of Chinese primary school students: A preliminary study. *Twelfth Congress of the European Society for Research in Mathematics Education (CERME12)*, *TWG08*(13). https://hal.archives-ouvertes.fr/hal-03745622

Prieto, G., & Nieto, E. (2014). Analysis of rater severity on written expression exam using Many Faceted Rasch Measurement. *Psicológica*, *35*(2), 385-397.

Thomson, S., Wernert, N., Buckley, S., Rodrigues, S., O'Grady, E., & Schmid, M. (2020). *TIMSS 2019 Australia. Volume II: School and classroom contexts for learning*. Australian Council for Educational Research. https://doi.org/10.37517/978-1-74286-615-4

Toffoli, S. F. L., de Andrade, D. F., & Bornia, A. C. (2016). Evaluation of open items using the many-facet Rasch model. *Journal of Applied Statistics, 43*(2), 299-316. https://doi.org/10.1080/02664763.2015.1049938