WHO BENEFITS FROM OPTIMIZING MATHEMATICALLY PRODUCTIVE STRUGGLE? INITIAL INSIGHTS FROM PROSPECTIVE ELEMENTARY TEACHERS' MATHEMATICAL JUSTIFICATIONS

Nitchada Kamlue, Laura R. Van Zoest
Western Michigan University

This exploratory qualitative study used 26 prospective elementary teachers' (PTs') mathematical justifications before and after they participated in a whole-class discussion of a high-leverage PT contribution to provide insight into who benefits from optimizing mathematically productive struggle. We used their own words to code and analyzed the 52 written responses, and identified five types of understanding: Robust, Substantial, Emerging, Preliminary, and No Evidence. We learned that, regardless of the PTs' initial type of understanding and whether they expressed struggle, 85% demonstrated progress after the discussion, and 69% evidenced an increase in their type of understanding. We conclude that optimizing productive struggle may be a particularly productive approach for developing PTs' flexible thinking.

INTRODUCTION

Productive struggle is becoming ubiquitous in mathematics education. Over the last five years, Scopus (Elsevier, n.d.) lists more than 400 peer-reviewed documents that contain both "productive struggle" and some variation of "math." There is also growing information about what teachers can do to support student struggle (e.g., Warshauer, 2015). The U.S. National Council of Teachers of Mathematics (NCTM, 2014) encourages teachers to "Support Productive Struggle in Learning Mathematics" (p. 48), as one of its eight teaching principles. Much of classroom instruction occurs as a whole class, yet Kamlue and Van Zoest (2025) found that even studies investigating productive struggle that have included data from whole-class mathematics discussions (e.g., Baker et al., 2024) did not focus on describing the learners' mathematical outcomes from engaging in productive struggle during those discussions. Other studies have identified the benefits of using students' thinking as a foundation for whole-class discussion (e.g., Bishop, 2021) and particular high-leverage student thinking on which to focus (e.g., Stockero et al., in press). In any classroom, there are likely to be students who are struggling (productively or unproductively) with a particular mathematical artifact (e.g., a task, an idea, a peer's thinking) and those who are not. This paper addresses the open question of whether each of those groups benefits from discussing a peer's thinking in an environment that uses what is known about optimizing productive struggle. We do this by investigating two research questions: (a) *What is the type of understanding reflected in the prospective elementary teachers' (PTs') mathematical justifications before and after they engage in a whole-class discussion*

2026. In H. Krzywacki, M. Hähkiöniemi, E. Haataja, M.S. Hannula & P. Hästö (Eds.). *Proceedings of the 49th Conference of the International Group for the Psychology of Mathematics Education: Research Reports*, pp. 1523-1530. PME.

of a PT's high-leverage contribution? and (b) *What can we learn about who benefits from optimizing mathematically productive struggle from the PTs' justifications and other indications of their mathematical progress?*

THEORETICAL PERSPECTIVES

We describe here five theoretical perspectives that undergird our study. First, Warshauer's (2015) framework for supporting productive struggle served as the starting point for optimizing mathematically productive struggle in our learning environment. This framework includes the use of *high cognitive demand tasks* (e.g., Stein et al., 2008). Second, Torres's Rights of the Learner (Kalinec-Craig, 2017), including the right to be confused, was a core component of the instructor's approach to teaching. It also informed our approach to elevating the learners' perspective in our data collection. Third, we take a sociocultural orientation to learning and thus identified the teaching practice of *building*, which "engages the whole class in a joint sense-making discussion focused on understanding the mathematics at the heart of a [high-leverage student mathematical contribution]" (Stockero et al., in press, Section 3, para. 1) as an important framework for our study. The four collections of teacher actions that make up what we refer to in this paper as a Building Episode were described by Stockero et al. (in press) as *Establish* the specific contribution that is to be the object of discussion, *Grapple Toss* that object to the class in way that positions them to make sense of it, *Conduct* a whole-class discussion that engages the class in making sense of the grapple-tossed object (the established contribution), and *Make Explicit* the underlying mathematics of the contribution that the class has come to better understand as a result of the discussion. Fourth, we define *mathematically productive struggle* as "a phenomenon in which learners expend and sustain efforts to make sense of a mathematical artifact while experiencing tension, and those efforts result in mathematical learning" (Kamlue & Van Zoest, 2025, p. 8) and use that definition as the foundation for determining the different ways the PTs experienced a Building Episode. Finally, we drew on Carpenter et al.'s (2003) three types of justifications to inform our thinking about using the PTs' responses to assess their understanding of the mathematics: *appeal to authority*—appealing to a rule, procedure, or someone's instruction without providing a mathematical reason; *justification by example*—using specific cases or examples to argue that something is true; and *generalizable arguments*—providing an argument that can be extended beyond the current situation.

METHODOLOGY

This report is part of a larger study investigating mathematical struggle during PTs' whole-class discussion of a high-leverage PT contribution. For more details about the methodology, see Kamlue and Van Zoest (2025). Our participants were 26 PTs enrolled in a course focused on developing children's abilities to reason numerically. We optimized opportunities for productive struggle through the choice of the course, the instructor, and the task, and by using a Building Episode to bound our analysis. Figure 1 shows the Focal Task that we used, the (incorrect) high-leverage PT

contribution that was the object of discussion for the approximately 15-minute Building Episode, and the underlying mathematics of the contribution that the Building Episode was intended to help the PTs to better understand. We collected information about the PTs' mathematical progress and struggle from five sources: (1) Focal Task Form 1 (Before the Building Episode); (2) In-the-Moment Reaction after the Grapple Toss; (3) Transcript of the recorded whole-class discussion; (4) Takeaway Reaction before Make Explicit; and (5) Focal Task Form 2 (After the Building Episode).

Figure 1: The Focal Task, the Focal Contribution, and the underlying mathematics

Focal Task	**(Incorrect) Focal PT Contribution**:	**Underlying Mathematics** of the Focal PT Contribution
Use words and/or pictures to justify your thoughts about: Which is Bigger? 6/7 or 7/8?	"6/7 is bigger than 7/8 because [they are each missing 1 piece and] sevenths are bigger than eighths"	*When comparing fractions to a larger benchmark, the smaller the amount missing from the benchmark, the larger the fraction.*

For the study reported here, we drew on two analyses. For the first analysis, we de-identified the 26 PTs' responses to the Focal Task before and after[1] the Building Episode and used *open coding* (Creswell & Poth, 2018) to capture the nuances of the PTs' understanding of the mathematics of the task. After we identified, described, and refined the *types of understanding* that represented the collection of 52 responses, we went back to the data and recoded them using the refined types (shown in Figure 2, column 1; see column 2 for representative examples from the data). Comparing the PTs' types of understanding before and after the Building Episode was a "high bar" measure of who benefited from optimizing mathematically productive struggle.

For the second analysis, we looked across all the data and used *open coding* with *in vivo* codes—using participants' exact words or phrases as codes (Creswell & Poth, 2018)—to capture: (a) evidence of mathematical struggle and (b) evidence of mathematical progress. The PTs were specifically asked to rate their confusion as a proxy for *mathematical struggle* on data sources 1, 2, and 5. We also searched for other types of evidence across all the data that reflected PTs' "efforts to make sense of a mathematical artifact while experiencing tension" (from our definition of Mathematically Productive Struggle). For example, a PT engaged in doing mathematics saying or writing, "I'm confused" or "I don't know," adding a question mark to a written statement, or pausing during an explanation.

We operationalized evidence of *mathematical progress* as PTs' verbal or written expressions indicating that they had made sense of something mathematical, regardless of whether that something was the task, the underlying mathematics of the contribution, or another mathematical idea introduced to the discussion. The PT's written responses on the Focal Task Form 1 served as an important reference point for identifying mathematical ideas they had made sense of as a result of their struggle. Looking at evidence of PTs' mathematical progress was a "low bar" measure of who benefited from optimizing mathematically productive struggle.

We considered PTs to have experienced (a) mathematically *productive* struggle if they demonstrated evidence of mathematical progress after they expressed their initial evidence of mathematical struggle, (b) mathematically *unproductive* struggle if they demonstrated evidence of mathematical struggle without evidence of mathematical progress, and (c) *no* mathematical struggle if they demonstrated no evidence of mathematical struggle.

RESULTS & DISCUSSION

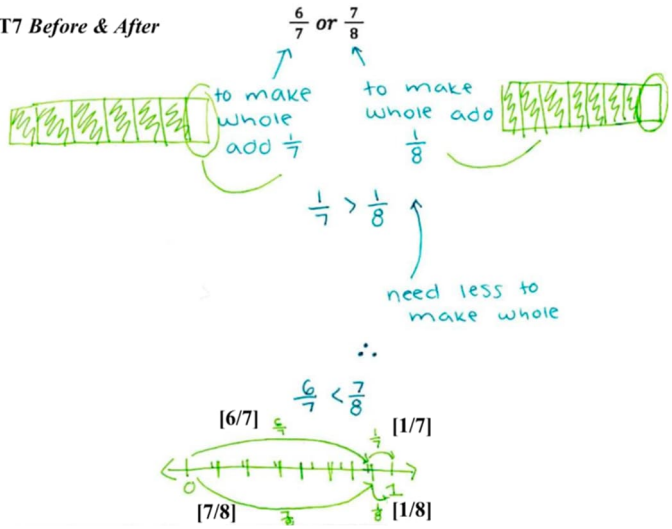
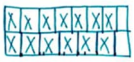
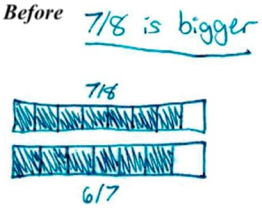
To address the open question of whether the focus on productive struggle called for by NCTM (2014) and others benefits both those who are struggling with a mathematical artifact and those who are not, we discuss (a) types of understanding reflected in PTs' mathematical justifications and (b) who benefited from participating in a whole-class discussion that optimized mathematically productive struggle.

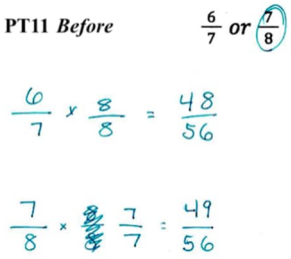
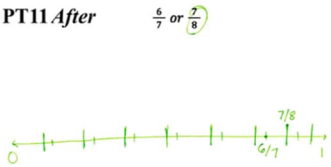
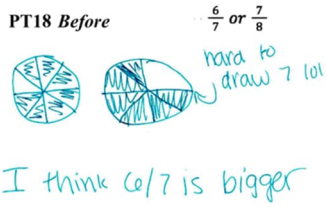
Types of Understanding Reflected in PTs' Mathematical Justifications

From our analysis of the set of 52 PT responses, we identified five types of understanding reflected in their mathematical justifications: *Robust*, *Substantial*, *Emerging*, *Preliminary*, and *No Evidence* (see Figure 2 for their definitions, illustrative examples, and numbers in each type). Two PTs demonstrated a robust understanding and eight a substantial understanding after the Building Episode, in contrast with one each before—reflecting a change from 8% to 38% of PTs who provided *generalizable arguments* (Carpenter et al., 2003). Another 35% of the PTs, those demonstrating an emerging understanding, had the beginnings of a generalizable argument after the Building Episode, in contrast to 12% before—an increase of 13%. Collectively, almost four times as many PTs were at or moving towards a generalizable argument after a 15-minute class discussion designed to optimize mathematically productive struggle (see Figure 3). For example, PT5's type of understanding changed from emerging (an accurate picture that had the potential to become a generalizable justification) to substantial when they added a justification that was generalizable (see Figure 2). PT6's type of understanding before the Building Episode was substantial—an accurate response and generalizable justification (see Figure 2) that required inferring that "1/8 / 1/7" referred to "the section that remains for each fraction" and that the 1/8 is smaller than the 1/7. Their type of understanding after the Building Episode was considered robust because they clarified their justification, writing, "Instead of thinking about the larger part, think about the smaller part that is left. So, in this case, look at the 1/7 and 1/8 that are left, and we know that 1/7 is larger than 1/8, so seeing that 1/7 is missing from the 6/7 we know that 6/7 is smaller than the 7/8. PT18's type of understanding changed from preliminary (an inaccurate picture without a beginning of a generalizable argument) to emerging when they provided a generalizable argument with an inaccurate answer: "I think that 6/7 is bigger because the smaller the piece missing, the bigger the fraction is." When PTs improved their type of understanding, like these examples, they met the "high bar" of benefiting from optimizing mathematically productive struggle.

The remaining 27% of PTs, coded Preliminary or No Evidence, were not yet moving towards a generalizable argument. However, there were some changes in the way they justified their answers. For example, PT11 shifted their justification from an *appeal to authority* (a procedure) to attempting to use a number line as a linear model of fractions to justify their answer (see Figure 2), a strategy that was shared during the whole-class discussion. Although this change did not move PT11 out of a preliminary understanding, they were considered to have made mathematical progress. PTs, like PT11, who broadened their thinking without advancing their type of understanding, were considered to have met the "low bar" of benefiting from optimizing mathematically productive struggle. PT25, on the other hand, changed their answer,

Figure 2: Type of understanding, PT examples, and number of PTs showing evidence of each type before and after the Building Episode.

Type of Understanding	PT Examples		Before	After
Responses That Provide, or Are Moving Toward Providing, a Generalizable Justification				
Robust ~ all aspects of the response are accurate, the justification is generalizable, and no inference is required.	**PT7** ***Before & After*** 		1 (3.9%)	2 (7.7%)
Substantial ~ all aspects of the response are accurate and the justification is generalizable, but some inference is required (for the justification or answer).	**PT6** ***Before***  "When looking at this problem, I would first think that 7/8 is larger than 6/7. I think this because the 1 section left (1/8 / 1/7) is smaller than the 1/7.	**PT5** ***After*** "I don't want to modify anything. The pieces in 6/7 are bigger than the pieces in 7/8. Therefore, 6/7 is further away from the whole than 7/8. That means 7/8 is bigger than 6/7."	1 (3.9%)	8 (30.8%)
Emerging ~ beginnings of a generalizable justification regardless of answer; or a generalizable justification with inaccurate or no answer.	**PT5** ***Before*** 	**PT18** ***After*** "I think that 6/7 is bigger because the smaller the piece missing, the bigger the fraction is."	3 (11.5%)	9 (34.6%)

Type of Understanding	PT Examples		Before	After
Responses That Are Not Yet Moving Toward a Generalizable Justification				
Preliminary ~ an appeal to authority (e.g., rule, algorithm, or someone's instruction) or only an attempt at picture(s) or diagram(s).	**PT11** ***Before*** 	**PT11** ***After*** 	18 (69.2%)	6 (23.1%)
	PT18 ***Before*** 	**PT25** ***After*** "I think my answer has changed because it was worked out as a whole group. So overall I think 7/8 is bigger than 6/7."		
No Evidence ~ no mathematics or only describes mathematics.	**PT24** ***Before*** "Initially just looking at them they looked different because one has bigger prices. The thing that's throwing me off is they have one less numerator and denominator."	**PT24** ***After*** "Now looking at it more and clearly seeing it with visuals now I know which one is bigger, and I was right. Knowing you have more of less units size is good is the big take away I had."	3 (11.5%)	1 (3.9%)

but did not provide evidence of mathematical progress in their *after* response (see Figure 2). PT24, in the only No Evidence *after* response, suggested increased confidence in their answer and potential attention to mathematical ideas related to unit size; however, their response also did not provide evidence of mathematical progress. The fact that seven PTs (8, 10, 11, 12, 24, 25, & 26; see Figure 3) did not move beyond a preliminary understanding is not surprising given that comparing two fractions whose numerator is one less than their denominator has been identified as particularly challenging for learners (Clarke & Roche, 2009).

Who Benefits From Optimizing Mathematically Productive Struggle

Figure 3 shows the changes in each PT's type of understanding and whether they demonstrated evidence of mathematical progress. Of the 26 PTs, 22 (85%) provided evidence of mathematical progress (meeting the "low bar" for benefitting from optimizing mathematically productive struggle), and 18 PTs' *after* justifications (69%) reflected an increase in their type of understanding (meeting the "high bar"). There was no evidence that any of the PTs' understanding was negatively affected by optimizing mathematically productive struggle. The PTs demonstrated growth in their mathematical understanding at their own pace. For example, while PT7 showed growth within a Robust type of understanding, PT9 shifted from No Evidence to Emerging. Additionally, although PT8 and PT9 showed No Evidence of understanding initially, *after* the Building Episode they demonstrated growth to Preliminary and Emerging

types of understanding, respectively. We also noticed that the entire Unproductive Struggle group was not yet moving towards a generalizable argument. In contrast, everyone in the No Mathematical Struggle group and 75% of the Mathematically Productive Struggle group were providing, or moving toward providing, a generalizable argument. The growth in both groups was similar, but 57% of the No Mathematical Struggle group was already providing, or moving toward providing, a generalizable argument in their *before* response to the Focal Task, in contrast with only 6% of the Mathematically Productive Struggle group.

Figure 3: PTs' struggle experiences and changes in PTs' mathematical understanding

		No Mathematical Struggle							Mathematically *Productive* Struggle																Unproductive		
	PT	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
Type of Understanding	Robust						P	P																			
	Substantial		P	P		P														P	P	P	P	P			
	Emerging	P			X					P				P	P	P	P	P	P								
	Preliminary								P		P	P	P													X	X
	No Evidence																								X		

Key: **Before** Building Episode **After** Building Episode P Progress X No Progress

Note: If the Before and After were the same type, only the After is visible (e.g., PT4, PT10, & PT25).

We look at PT7, the only PT who had a robust understanding before the Building Episode and who did not express any struggle, for insight into whether such learners might benefit from optimizing mathematically productive struggle. (See Figure 2 for PT7's responses.) Although PT7's *before* justification was already a *generalizable argument*, after the Building Episode, they added both an area model and a linear model (number line) to their justification. This suggests that they also benefited from the optimizing mathematically productive struggle experience by considering ways multiple representations could be used to communicate the fraction comparison.

CONCLUSION

A whole-class discussion designed to optimize productive struggle can benefit both those who actually struggle productively and those who do not. We found that the majority of the prospective elementary teachers' (PTs') mathematical justifications after they engaged in a whole-class discussion of a PT's high-leverage contribution demonstrated growth in their understanding—regardless of whether they struggled themselves or their initial understanding. Of particular note is the fact that even the one PT who had a robust understanding of mathematics before the discussion broadened their response to the task in a way that represented a more flexible understanding. This research suggests that optimizing productive struggle by engaging PTs in whole-class discussions in which they collectively make sense of a peer's contribution may be a particularly productive approach for teacher education.

Endnote

[1]PT6 provided a revision of their focal task justification in the Takeaway Reaction at the end of the Building Episode, but did not repeat this information on Focal Task Form 2. Because omitting the revision would have produced an inaccurate

representation of PT6's understanding, we used the revised justification to evaluate their type of understanding after the Building Episode. For all other PTs, Focal Task Form 2 was consistent with other evidence of growth we had for them.

Acknowledgements

The first author gratefully received a David P. Y. Burns Mathematics Finishing Scholarship, Western Michigan University, to support her work on this paper as part of her dissertation completed under the direction of the second author.

References

Baker, K., Clark, M., & Gallagher, D. M. (2024). Together in a productive struggle: Unpacking student and teacher productive struggle in mathematics. *The Electronic Journal for Research in Science & Mathematics Education*, *27*(4), 107-117.

Bishop, J. P. (2021). Responsiveness and intellectual work: Features of mathematics classroom discourse related to student achievement. *Journal of the Learning Sciences 30*(3), 466-508. https://doi.org/10.1080/10508406.2021.1922413

Carpenter, T. P., Franke, M. L., & Levi, L. (2003). Thinking mathematically: Integrating arithmetic and algebra in elementary school. Heinemann.

Clarke, D. M., & Roche, A. (2009). Students' fraction comparison strategies as a window into robust understanding and possible pointers for instruction. *Ed. St. in Math., 72*, 127-138.

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). Sage publications.

Elsevier. (n.d.). *Scopus* [Database]. Elsevier. https://www.scopus.com

Kalinec-Craig, C. A. (2017). The rights of the learner: A framework for promoting equity through formative assessment in mathematics education. *Democracy and Education*, *25*(2), 1-11.

Kamlue, N., & Van Zoest, L. R. (2025). *An Exploratory Study of Mathematically Productive Struggle in Prospective Elementary Teachers' Whole-Class Discussion of a High-Leverage Contribution.* [Manuscript submitted for publication]. Department of Mathematics, Western Michigan University.

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

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.

Stockero, S. L., Leatham, K. R., & Peterson, B. E., Van Zoest, L. R. (in press). A conceptualization of the teaching practice of building. *Journal of Mathematics Teacher Education.*

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