Argumentation, Justification, and Proof in Middle Grades: A Rose by Any Other Name

Eric Knuth, Orit Zaslavsky, and Hangil Kim

Introduction

We were tasked by the editors to provide a synthesis of the three chapters, highlighting, in particular, the similarities and differences among the chapters as well as the consequences associated with analyzing the classroom data from the three different perspectives of argumentation, justification, and proof. In our discussion, we also highlight the consequences associated with the analytic framework that each set of authors brought to bear on the data and suggest that these latter consequences may contribute more to the differences among the chapters than the differences in perspective. In addition, we also offer an alternative definition for the activities associated with argumentation, justification, and proof as we think it may serve as a more comprehensive way for describing and studying such activities.

The chapter authors were provided definitions of the three focal constructs, definitions that the editors acknowledge are one of many ways that each particular construct might be defined. Indeed, as Hanna (2020) claimed: "Argumentation, reasoning, and proof are concepts with ill-defined boundaries" (p. 561). Likewise, we would add to that list of concepts, *justification*, as it too falls into the category of concepts with ill-defined boundaries. Nevertheless, in our discussion we used the following definitions of each construct as stated by each of the chapter authors:

E. Knuth (✉) · H. Kim
University of Texas at Austin, Austin, TX, USA
e-mail: eric.knuth@austin.utexas.edu; hangil_kim@utexas.edu

O. Zaslavsky
New York University, New York, NY, USA
e-mail: orit.zaslavsky@nyu.edu

Argumentation as "the process of making mathematical claims and providing evidence to support them" (see "Exploring Collective Argumentation in the Mathematics Classroom" in the chapter by Gomez Marchant, Jones, and Tanck, this volume).

Justification as "the process of supporting your mathematical claims and choices when solving problems or explaining why your claim or answer makes sense" (Bieda & Staples, 2020)" (see the introductory paragraph in the chapter by Lesseig and Lepak, this volume).

Proof as an argument that consists of six characteristics, with arguably the most critical characteristic being that the argument "does not admit possible rebuttals" (see "Introduction" in the chapter by Yopp, Ely, Adams, and Nielsen, this volume).

Based on these working definitions, we view proof as a special case of argumentation and view justification as the second "half" of argumentation (i.e., providing evidence to support a claim). The relationship among these three constructs serves to inform our discussion of the three chapters. In our discussion of the chapters, we also consider aspects of practice not necessarily captured with argumentation, justification, or proof, namely, the practices of developing and exploring conjectures. We refer to the collective set of practices as *proving-related activities* (Knuth et al., 2019) and use this expression throughout our discussion in this chapter.

The Task and Its Possibilities

In discussing the three chapters, we first consider the potential of the task to engage students in proving-related activities (cf. Arbaugh et al., 2019) and the instructional goals associated with the Number Trick task as each of the chapter authors based aspects of their analyses on assumptions about the task and on its implementation.[1] In their chapter, Lesseig and Lepak introduced the Mathematics Task Framework (Stein et al., 1996) as a means of capturing "a task from its original form [as written] through different phases of set up and implementation, culminating in student learning" (see "Method" in the chapter by Lesseig and Lepak, this volume). We adopt this particular lens as well to frame our discussion as it offers a way to portray the possibilities afforded by the written task, the potential instructional goals associated with the task, and the opportunities, both taken and missed, by the teacher to meaningfully engage students as the task was set up and enacted.

---

[1] We are uniquely positioned to discuss the potential outcomes and instructional goals of the Number Trick task as the first author used this same task in a project which examined middle school students' justifying and proving competencies (e.g., Knuth & Sutherland, 2004; Knuth et al., 2009), and much of our discussion is based on this prior work.

Potential Proving-Related Activities

In the first part of the task, students are asked whether Jessie's two answers will always be equal to each other for any number between 1 and 10 and then asked to explain their reasoning. Responses that might be anticipated from students include (i) testing several, but not all, numbers between 1 and 10 and concluding that the two answers *will* always be equal based on the results of their computations; (ii) testing every number between 1 and 10 and concluding that the two answers *are* always equal; or (iii) testing several numbers between 1 and 10, gaining a critical insight from the computational process (i.e., an insight related to the distributive property), and concluding that the two answers *will* always be equal based on that insight. In all three cases, students' explanations involve reasoning with examples: in the first case, an explanation based on a limited set of empirical evidence; in the second case, an explanation based on exhausting the domain of the claim; and in the third case, an explanation based on observing generality in the computational process. In the last two cases, one might consider the explanations to constitute proof: a proof-by-exhaustion and a generic proof (Zaslavsky, 2018) or an algebraic proof (i.e., $2(x + 4) = 2x + 8$).

In the second part of the task, students are asked, with respect to their explanation to the first part, whether their explanation shows that the two answers will always be equal to each other for any number (not just between 1 and 10) and to explain their response. Anticipated student responses might include (i) concluding that the answers will always be equal for any number based on their work from the first part of the task (either their computations with several numbers or their computations with every number); (ii) testing several additional numbers (beyond the domain from the first part of the task) and concluding that the two answers *will* always be equal for any number based on these additional computations; (iii) concluding that they cannot be sure because they only tested a limited set of numbers (in the first part) and cannot test every possible number; or (iv) concluding that the two answers will always be equal based on the critical insight related to the distributive property. Similar to the first part of the task, students' explanations again involve reasoning with examples: the first two cases both rely on a limited set of empirical evidence, and the fourth case relies on noticing the generality of the computational process. With respect to whether the explanations constitute proof, as above, only the fourth case might be considered as such.

In considering the three chapters in light of the preceding discussion, two of the three chapters do focus, to varying degrees, on aspects of the task and possible student responses. Yopp et al. noted that the task offers the possibility for students to recognize "that a proof of a general claim needs more than just an empirical check" (see "Reframing the Definition of Proof for the Current Context" in the chapter by Yopp et al., this volume), to engage in proof-by-exhaustion (for the first part of the task) and to gain a conceptual insight regarding regularities in the computational process of testing several numbers. Although Lesseig and Lepak also noted this last possibility, it was only in the context of the second part of the task as they claimed

that the computational process students used for the first part of the task only served the dual purpose of leading "students to an answer that the Number Trick worked (the claim) and provided the explanation to support their claim" (see "Justification Process" in the chapter by Lesseig and Lepak, this volume). Gomez, Marchant Jones, and Tanck, however, did not discuss the nature of the task and its possibilities as their analytic framework focuses solely on the implementation of the task.

Potential Instructional Goals

In considering the task setup and its implementation, the three anticipated student responses for the first part of the task provide various instructional opportunities for a teacher to engage students in proving-related activities. And, of course, these instructional opportunities are also enabled (and constrained) by a teacher's particular instructional goals. A teacher might initially ask students who tested several examples how they know the two answers will be the same for the numbers they did not test, and in the classroom conversation that might follow, other students might offer critiques of the responses. A teacher's instructional goal in this case might be to highlight the limitation of examples as a means of establishing a general result. Alternatively, a teacher might facilitate a conversation between students who tested several examples and those who tested the entire domain, asking the students to compare (and critique) the respective strategies. In this case, such a conversation might serve an instructional goal of contrasting the limitation of examples as a means of establishing a general result with their power to prove a general result (as well as to foreshadow what is to come in the second part of the task). Moreover, a teacher might also have a goal of gaining insight about the rationale underlying students' decisions to test every number in the domain: on the one hand, it might be that students did so because they assumed that the task (or teacher) required it, and on the other hand, it might be that students viewed the strategy as a means of gaining absolute certainty about the truth of the claim (a strategy that, as Yopp and his colleagues might say, does "not admit rebuttals"). Finally, for those students who gained a critical insight from noticing regularities in the computational process, having these students share their thinking might serve as a natural transition into the second part of the task (or a teacher might wait until the class has thought about the second part of the task before asking these students to share their thinking).

As the class moves on to the second part of the task, a teacher might first ask for responses from those students who concluded the answers will always be the same based on their computational results from the first part of the task (or based on their testing of additional examples). In the class discussion that follows, other students might suggest that it is impossible to test every number, so one cannot be absolutely sure that the two answers will always be the same for every possible number—an opportunity to highlight the limitations of examples as a means to establish a general result, beyond any doubt (or without admitting rebuttals). Finally, in asking students if there is any way to know for certain whether the two answers will always

be the same for any number, students who gained the critical insight (or at least searched for one) could offer their explanations—an opportunity to discuss the idea of a general argument (versus an argument based on empirical evidence).

In examining how the task was set up and implemented in Mr. MC's classroom, each set of chapter authors highlighted different aspects of the classroom interactions as students engaged with the task. Gomez Marchant and his colleagues do not discuss the way in which the task was set up but rather used the Teacher's Support for Collective Argumentation [TSCA] framework (Conner et al., 2014) to document teacher and student contributions in the argumentation process and, importantly, highlighted the teacher moves in that process. The TSCA framework provided a means for examining the nature of the teacher's contributions to the process, whether adding contributions himself or eliciting contributions from students, and documents the critical role the teacher played in supporting the classroom's collective argumentation. Interestingly, and somewhat in contrast to the Lesseig and Lepak chapter, Gomez Marchant et al. do not comment about any of the opportunities Mr. MC may have missed to engage students in thinking more deeply about the task. For example, Lesseig and Lepak noted that Mr. MC pushed students to compute every number for the first part of the task ("Did you check them all?") and "seemed to discourage a general line of reasoning in favor of a justification for 1–10 early in the process" (see "Task as Implemented" in the chapter by Lesseig and Lepak, this volume). In this case, by pushing students to compute every number between 1 and 10, the teacher removed the possibility, at that moment, for discussing different student responses (e.g., testing only a few numbers, noticing the underlying structure). It is also not clear whether Mr. MC intended for students to test every number in the domain (part 1) with proof-by-exhaustion in mind. If a goal was to use the task as context for introducing the distributive property through a discussion of the computational regularities students may have noticed, and to then use a generic example to illustrate the structure, this opportunity was missed.

Framing Matters

Across the three chapters, the analytic framework applied to the data, in large part, seemed to make the biggest difference in what the chapter authors reported. Gomez Marchant et al.'s use of the TSCA framework (Conner et al., 2014) resulted in a primary focus on documenting both student and teacher moves in creating an argument, whereas Lesseig and Lepak's use of the Mathematics Task Framework (Stein et al., 1996) resulted in a primary focus on the teacher's instructional practices and the resulting opportunities for students to learn from their engagement in the justification process. In their chapter, Yopp and his colleagues focused less on the actual classroom implementation of the task and more on the "product" of students' engagement in the proving-related activities. In particular, they analyzed the student work relative to six criteria for what constitutes a proof and characterized the responses students produced in light of these criteria. An important distinction that

Yopp et al. noted relates to the critical insight about the distributive property: they distinguish between students who relied solely on empirical evidence and those students who either searched for a "conceptual" insight (but did not find) or determined the conceptual insight. The search for a conceptual insight is important to note as that suggests that these students do recognize the limitation of empirical evidence.

In the end, it is not surprising that each set of chapter authors, given their respective constructs and definitions (i.e., argumentation, justification, or proof) and particular analytic framework, detailed different aspects of the classroom participants' interactions (and products). We also think it is worth noting that none of the chapter authors discussed in any detail the various possible outcomes and instructional goals associated with the Number Trick task as written nor the potential opportunities for deepening students' understanding of argumentation, justification, and proof. From our perspective, the Number Trick task offered several opportunities for students to learn important ideas underlying proving-related activities including the limitation of examples as a means of proof as well as their power as a means of proof (in the case of proof-by-exhaustion or in generic proving), the role of examples can play in noticing regularity or gaining conceptual insights (from computations with examples), and what constitutes a general argument (proof). Yet, in their analyses of the classroom episode, the chapter authors did not discuss in any depth such opportunities (whether taken or missed by the teacher). For example, the authors all noted to varying degrees the fact that Mr. MC pushed his students to test every case between 1 and 10, yet they did not comment that, as a result, Mr. MC was taking away the opportunity to engage the class in a conversation about the role of empirical evidence (its limitations as well as its power).

Capturing the Breadth of Proving-Related Activity

In considering the perspectives of the three chapters—argumentation, justification, and proof—we are reminded of a quote from William Shakespeare: "A rose by any other name would smell as sweet." In this case, whether one uses argumentation, justification, or proof, it seems that the descriptive name of the activity is of less importance than the actual nature of the activity in which students engage. Stylianides (2009), for example, defined reasoning-and-proving:

> to describe the overarching activity that encompasses the following major activities that are frequently involved in the process of making sense of and establishing mathematical knowledge: identifying patterns, making conjectures, providing non-proof arguments, and providing proofs. The choice of a hyphenated term to encompass these four activities reflects my intention to view the activities in an integral way. (pp. 258–259)

We agree with Stylianides in the need to describe the overarching activity but feel his definition is not adequately comprehensive in that it excludes some related activities. As we mention in the introduction, we prefer the expression *proving-related*

activities as we view it as encompassing aspects of practices related to argumentation, justification, and proof as well as aspects of practices not necessarily captured by these three practices (Ellis et al., 2019). In particular, we characterize proving-related activities as including the development of conjectures, exploration of conjectures, justification of conjectures (including proof of conjectures), and refutation of conjectures. Inherent in our definition is also explicit attention to the role examples play in these activities as we view the time spent thinking about and analyzing examples as playing a foundational and essential role in the development, exploration, and understanding of conjectures, as well as in subsequent attempts to develop proofs of those conjectures.

A focus on argumentation, justification, or proof (as defined and applied by the authors), from our perspective, does not adequately capture all of the critical aspects of proving-related activity, including the development of conjectures as well as the role of examples-based reasoning. As the Number Trick task highlights, examples played a major role in the students' activities, from providing initial conviction about the claim's truth to serving as a means of justification for the claim's truth to providing insight based on regularities observed in the computations with examples.

One final consideration for a more comprehensive way to describe the instructional episode that was the focus of the three chapters relates to our collective work with teachers. If we want the activities associated with argumentation, justification, and proof to play a more central role in middle school classrooms, it may be more powerful and instructive to focus on all the activities involved in such practices. It is easy to get lost in trying to categorize teachers' practices or students' activities as related to argumentation, justification, or proof, when in the end what perhaps really matters is teachers' efforts to meaningfully engage students in the proving-related activities of developing, exploring, justifying (including proving), and refuting conjectures.

References

Arbaugh, F., Smith, M., Boyle, J., Stylianides, G. J., & Steele, M. (2019). *We reason and we prove for all mathematics: Building students' critical thinking, grades 6-12*. Corwin, a Sage Company.

Bieda, K., & Staples, M. (2020). Justification as an equity practice. *Mathematics Teacher: Learning and Teaching PK-12, 113*(2), 102–108.

Ellis, A., Ozgur, Z., Vinsonhaler, R., Dogan, M., Carolan, T., Lockwood, E., Lynch, A., Sabouri, P., Knuth, E., & Zaslavsky, O. (2019). Student thinking with examples: The criteria-affordances-purposes-strategies framework. *The Journal of Mathematical Behavior, 53*, 263–283.

Hanna, G. (2020). Mathematical proof, argumentation, and reasoning. In S. Lerman (Ed.), *Encyclopedia of mathematics education* (pp. 561–566). Springer.

Knuth, E., Choppin, J., & Bieda, K. (2009). Middle school students' production of mathematical justifications. In D. Stylianou, M. Blanton, & E. Knuth (Eds.), *Teaching and learning proof across the grades: A K-16 perspective* (pp. 153–170). Routledge.

Knuth, E. & Sutherland, J. (2004). Student understanding of generality. In D. McDougall & J. Ross (Eds.), *Proceedings of the Twenty-sixth annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 561–567). Toronto, Ontario: OISE/UT.

Knuth, E., Zaslavsky, O., & Ellis, A. (2019). The role and use of examples in learning to prove. *The Journal of Mathematical Behavior, 53*, 256–262.

Stein, M. K., Grover, B., & Henningsen, M. (1996). Building student capacity for mathematical thinking and reasoning: An analysis of mathematical tasks used in reform classrooms. *American Educational Research Journal, 33*(2), 455–488.

Stylianides, G. (2009). Reasoning-and-proving in school mathematics textbooks. *Mathematical Thinking and Learning, 11*(4), 258–288.

Zaslavsky, O. (2018). Genericity, conviction, and conventions: Examples that prove and examples that don't prove. In A. Stylianides & G. Harel (Eds.), *ICME-13 monograph: Advances in mathematics education research on proof and proving* (pp. 283–298). Springer.

Eric Knuth is a professor of STEM education in the College of Education at the University of Texas at Austin. His program of research focuses on mathematical practices related to algebraic reasoning and to mathematical argumentation (justifying and proving). A primary goal of his work is to develop sufficiently detailed accounts of learning and instruction in classroom contexts to guide the development of instructional practices, curricular materials, and professional development aimed at facilitating students' mathematical development.

Orit Zaslavsky is a professor in the Steinhardt School of Culture, Education, and Human Development and the Director of the Mathematics Education Program at New York University. Zaslavsky's recent work focuses on the interplay between mathematical examples, definitions, and proof. She explores how students think with and through examples when they are engaged in mathematical activities (e.g., proving) and how mathematics teachers use and think about examples when they plan their teaching and when they actually teach.

Hangil Kim is a doctoral student in STEM Education at the University of Texas at Austin. Prior to his doctoral studies, he taught mathematics for 3 years at a middle school and 3.5 years at a high school in South Korea. His research interests include a focus on proof and reasoning, the design of classroom opportunities to engage students in proving-related activities, and the design of teachers' instructional practices that facilitate students' engagement in proving-related activities.