TEACHERS' IDENTIFICATION OF CHILDREN'S UPPER AND LOWER BOUNDS REASONING

Carolyn A. Maher
Rutgers University
carolyn.maher@gse.rutgers.edu

Marjory F. Palius
Rutgers University
marjory.palius@gse.rutgers.edu

James A. Maher
Rutgers University

Robert Sigley
Rutgers University
robert.sigley@gse.rutgers.edu

This research investigates the question of what growth, if any, is shown by teachers in identifying the components of children's reasoning using an upper and lower bounds argument for a fraction task. Specifically, it reports on assessment outcomes from design-based research in teacher education that measures teachers' identification of children's reasoning from studying videos. We describe the nature of the instructional intervention as well as the video-based assessment used as a pre and post measures for identifying children's mathematical reasoning, and report on the nature of teacher growth in recognizing components of children's arguments.

Keywords: Reasoning and Proof; Teacher Education–Inservice/Professional Development; Rational Numbers; Design Experiments

Introduction

The research presented here comes from an ongoing, interdisciplinary research and development project[1] at a large public university. Work includes the development of a digital repository that provides open access to a seminal video collection of children's mathematical reasoning that accumulated through a quarter century of research on the development of mathematical thinking and reasoning in students.[2] Videos from the repository have been used to conduct design research in teacher education, specifically for the purpose of examining how the opportunity to study videos may help teachers augment their abilities to recognize mathematical reasoning as it emerges from children's explanations and justifications of their problem solving. Instructional interventions for teachers were created for implementation in courses or workshops, typically based on one of two models (Palius & Maher, 2011). We report here on a different kind of intervention model that was created specifically for implementation in the context of online learning with digital resources.

Theoretical Perspective

Learning occurs in complex contexts and it is important that it be studied in the way it naturally occurs (Brown, 1992; Greeno & MAP, 1998; Spiro, Feltovich, Jacobson, & Coulston, 1992). However, teachers and those preparing to be teachers do not ordinarily have the opportunity to study in detail the learning of individual students in classrooms. Collections of video offer a rich source of data for careful analysis and reflection on children's learning. Choosing subsets of videos from large collections can provide a rich resource for addressing particular research questions. Our work and the work of others have demonstrated that there is much to gain from studying episodes of children's learning from videos (Cobb, Wood, & Yackel, 1990; Maher & Davis, 1995; Fenemma, Carpenter, Franke, Levi, Jacobs, & Empsom, 1996; Tirosh, 2000). Further, video offers an excellent medium for teachers' development of what Bransford et al. (2006) refer to as "adaptive expertise," that is, an ability to spontaneously and flexibly identify, critically evaluate, and respond in appropriate ways to instances of children's learning. It is from this perspective that our study was designed.

Yackel and Hanna (2003) discuss the importance of reasoning and proof in mathematics learning and their functions of verification, explanation, and communication. They point to the need for mathematics educators to be able to support students' development along the continuum from reasoning, explaining,

Van Zoest, L. R., Lo, J.-J., & Kratky, J. L. (Eds.). (2012). *Proceedings of the 34th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education.* Kalamazoo, MI: Western Michigan University.

and justifying towards articulation of formal proof, as well as to the need for teachers to create a classroom atmosphere that support such development (Yackel & Hanna, 2003). Mathematics teacher education, therefore, is faced with the challenge of helping teachers to attend to emerging forms of reasoning as children express justifications using their own language. Making use of episodes and transcripts of video data of children's reasoning from a major collection, we sought to investigate whether teachers could build the mathematical knowledge for recognizing components of children's reasoning. Specifically, the question that guided our research was whether and to what extent teachers successfully identified components of children's reasoning using an upper and lower bounds argument for a fraction task.

Methodology

As part of the design research in teacher education, three of the authors developed a new, online course in mathematics education, entitled *Critical Thinking and Reasoning*, to be taken as an elective by graduate students. Its purpose was to focus teachers' attention to how children reason about fraction ideas through study of videos children's reasoning, while engaged in problem solving with fraction tasks (Yankelewitz, Mueller, & Maher, 2010). Research literature connected to the video content was assigned as readings to comprise course units around which online discussions were focused. As a component of the design research, we examined teachers' attention to children's reasoning before and after the intervention. For this report, we investigate the nature of teacher growth in identifying upper and lower bounds reasoning in children from videos.

The first implementation of the course was during a semester with 12 students participating in the research. The second iteration was done as a four-week summer session course with 10 students participating in the research. Both courses contained a unit that focused specifically on children's mathematical reasoning about the fractions task in the video assessment. Specifically, students were assigned to study two videos, *Fractions, Grade 4, Clip 1 of 4: David's upper and lower bound argument* (http://hdl.rutgers.edu/1782.1/rucore00000001201.Video.000054465) and *Fractions, Grade 4, Clip 4 of 4: Designing a new rod set* (http://hdl.rutgers.edu/1782.1/rucore00000001201.Video.000054751). The reading assignment from the unit was a book chapter that discussed children's mathematical exploration that leads toward proof-like reasoning, which included the example of David's upper and lower bounds argument (Maher & Davis, 1995). The prompt for group online discussions was open-ended and suggested that attention be paid to forms of children's arguments and the evidence they provide, as well as consideration of what may be evidence of understanding or evidence of obstacles to the children's understanding of the mathematics. Students were assigned to small groups for engaging in online discussions about the videos they were viewing and the related literature.

Consistent with methodology of the larger research project, participants were administered pre and post-tests to measure change from before to after the intervention. We focus here on a video-based assessment for identifying children's mathematical reasoning on a particular task in the fractions strand. The assessment video includes footage from research conducted in an after-school enrichment program for 6^{th} graders in an urban community, where children engaged in many of the same tasks that were explored by children in the 4^{th} grade classroom study (Maher, Mueller, & Yankelewitz, 2009). It contained short clips of children working in groups on a task to find a Cuisenaire rod in the set that could be given the number name one-half when the blue rod has been given the number name one. It also contained short clips of children explaining their solution ideas with rod models as justification to the whole class (Maher, Mueller, & Palius, 2010).

The children in the assessment video offered various explanations for why they found that there is no rod in the set that can be called one half when the blue rod is called one. Some of the explanations took the form of reasoning by cases; however, one of the arguments took the form of reasoning by upper and lower bounds (Yankelewitz, Mueller, & Maher, 2010). More than one child's discourse contributed to the articulation of this argument form, which, along with the mathematical sophistication of the argument, made it particularly interesting as focal point of analysis after coding the assessment data. That is, we were curious about the extent to which teachers would recognize that children were expressing in their own language that the solution for half of Blue is bounded by the Yellow and Purple rods, with Yellow being the least upper bound and Purple being the greatest lower bound (i.e., that there is no rod in between them).

A highly detailed rubric was developed by our research team in order to code the data by the components of the arguments that were articulated by the children in the assessment video. The assessment prompted study participants to describe as completely as they can the reasoning that the children put forth, whether each argument offered by children is convincing, and why or why not are they convinced. Participants were provided with a transcript for the video and were not restricted in the amount of time spent working on the assessment. The assessment prompt also informed participants that their responses would be evaluated by the following criteria: recognition of children's arguments, their assessment of the validity or not of children's reasoning, evidence to support their claims, and whether the warrants they give are partial or complete.

Two researchers scored assessment data with 90.4% inter-rater reliability. For the upper and lower bounds argument, there were four components of the children's reasoning that could combine in three different ways to be a complete argument (a, b, and c; a, b, and d; or a, b, c, and d):

a. The Yellow rod is (1/2 of one White rod) longer than half of Blue; (AND)
b. Purple is (1/2 of one White rod) shorter than half of Blue; (AND)
c. There is no rod with a length that is between Yellow and Purple; (OR)
d. The White rod is the shortest rod and the difference between the Yellow rod and the Purple rod is one White rod.

Participant responses that did not mention any of the above components or that mentioned only one or two of them were deemed to be incomplete. The coded data were analyzed quantitatively.

Results

Analysis of the video assessment data yielded the following results with regard to the upper and lower bounds argument. Tables 1a, 1b, and 1c describe the distributions of pre-assessment argument components, showing results for the two classes combined and then disaggregated by the two implementations of the course. In Table 1a, we note that 13 of the 22 students in the combined courses provided an incomplete argument description in the pre-assessment, while 8 of these 13 students provided none of the 3 essential argument components (a, b, and c or d) of a complete upper and lower bounds argument. A total of 11 out of 13 excluded argument component a; 12 out of 13 excluded argument component b; and 10 out of 13 excluded either argument component c or d. Table 1b shows that 8 of 12 students in the intervention provided an incomplete argument description in the pre-assessment; 5 of these 8 students provided none of the 3 essential components (a, b, c or d) of a complete argument description. A total of 3 out of 8 excluded argument component a; 7 out of 8 excluded argument component b; and 6 out of 8 excluded either argument component c or d. Table 1c shows that 5 of 10 students in the summer course intervention provided an incomplete argument description in the pre-assessment; 3 of these 5 students provided none of the three essential components (a, b, c or d) of a complete argument description. A total of 4 out of 8 excluded argument component a; 5 out of 5 excluded component b; and 4 out of 5 excluded either component c or d.

In summary, the pre-assessment results indicate that 59% of the students in the two courses did not provide a complete upper and lower bounds argument description on the pre-assessment. Of the students with an incomplete argument description, over 75% from the two combined courses failed to describe each of the three essential upper/lower bound argument components.

Van Zoest, L. R., Lo, J.-J., & Kratky, J. L. (Eds.). (2012). *Proceedings of the 34th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education.* Kalamazoo, MI: Western Michigan University.

Table 1a: Distribution of Pre-Assessment Argument Components: Two Courses Combined

Students with Incomplete Argument			Students with Complete Argument		
Components	Count	Frequency	Components	Count	Frequency
None	8	0.6154	a, b, c	4	0.4444
a	1	0.0769	a, b, d	2	0.2222
c	1	0.0769	a, b, c, d	3	0.3333
d	2	0.1538			
a, b	1	0.0769			
Total	13	1.0000	Total	9	1.0000

Table 1b: Distribution of Pre-Assessment Argument Components: Semester Course

Students with Incomplete Argument			Students with Complete Argument		
Components	Count	Frequency	Components	Count	Frequency
None	5	0.625	a, b, c	2	0.25
c	1	0.125	a, b, d	1	0.50
d	1	0.125	a, b, c, d	1	0.25
a, b	1	0.125			
Total	8	1.000	Total	4	1.00

Table 1c: Distribution of Pre-Assessment Argument Components: Summer Course

Students with Incomplete Argument			Students with Complete Argument		
Components	Count	Frequency	Components	Count	Frequency
None	3	0.6	a, b, c	2	0.4
a	1	0.2	a, b, d	1	0.2
d	1	0.2	a, b, c, d	2	0.4
Total	5	1.0	Total	5	1.0

Tables 2a, 2b, and 2c describe the distributions of post-assessment argument components, showing results for the two classes combined and then disaggregated by the two implementations of the course. In Table 2a, we note that of the 10 of the 22 students in the combined courses provided an incomplete argument description in the post-assessment, while only 1 of these 10 students provided none of the three essential argument components (a, b, and c or d) of a complete upper and lower bounds argument. A total of 4 out of 10 excluded argument component a; 4 out of 10 excluded component b; and 7 out of 10 excluded either component c or d. Table 2b shows that 6 of 12 students in the intervention provided an incomplete argument description in the post-assessment. Of these 6 students, at least one the three essential components (a, b, c or d) were provided. Three of the 6 students excluded argument component a; none excluded component b; and 5 out of 6 excluded either argument component c or d. Table 2c indicates that 4 of 10 students in the summer course provided an incomplete argument description in the post-assessment, 1 of these 4 students provided none of the three essential components a, b, c or d of a complete argument description. A total of 2 out of 4 excluded argument component a, 3 out of 4 excluded argument component b, and 2 out of 4 excluded either argument component c or d.

Van Zoest, L. R., Lo, J.-J., & Kratky, J. L. (Eds.). (2012). *Proceedings of the 34th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education.* Kalamazoo, MI: Western Michigan University.

Table 2a: Distribution of Post-Assessment Argument Components: Two Courses Combined

Students with Incomplete Argument			Students with Complete Argument		
Components	Count	Frequency	Components	Count	Frequency
None	1	0.1	a, b, c	4	0.3333
b	2	0.2	a, b, d	3	0.2500
d	1	0.1	a, b, c, d	5	0.4166
a, b	4	0.4			
a, d	1	0.1			
a, d	1	0.1			
Total	10	1.0	Total	12	1.0000

Table 2b: Distribution of Post-Assessment Argument Components: Semester Course

Students with Incomplete Argument			Students with Complete Argument		
Components	Count	Frequency	Components	Count	Frequency
b	2	0.3333	a, b, c	2	0.3333
a, b	3	0.5000	a, b, d	2	0.3333
b, d	1	0.1667	a, b, c, d	2	0.3333
Total	6	1.0000	Total	6	1.0000

Table 2c: Distribution of Post-Assessment Argument Components: Summer Course

Students with Incomplete Argument			Students with Complete Argument		
Components	Count	Frequency	Components	Count	Frequency
None	1	0.25	a, b, c	2	0.3333
d	1	0.25	a, b, d	1	0.1667
a, b	1	0.25	a, b, c, d	3	0.5000
a, d	1	0.25			
Total	4	1.00	Total	6	1.0000

In summary, the post-assessment results indicate that 45.5% of the students in the two courses combined were not able to provide a complete upper and lower bounds argument description, compared to 59% on the pre-assessment. Of the students with an incomplete argument description on the post-assessment, 40% failed to describe each of the components a and b, and 70% failed to describe component c or d. This is in contrast to over 75% who failed to describe each of the three argument components on the pre-assessment.

Table 3 classifies the pre-assessment argument descriptions into three categories: (1) a Complete Argument description containing components a, b, and c or d; (2) a No Components description which lacks all three essential argument components; and (3) a Partial Argument description which contains at least one essential argument component but lacks all three. The respective frequencies for the two combined courses are: 40.9% Complete Argument, 36.4% No Argument Components, and 22.7% Partial Argument.

Table 3: Upper-Lower Bound Pre-Assessment Argument Frequencies

Pre-Assessment Argument Components	Combined Courses		Semester Course		Summer Course	
	No.	Freq.	No.	Freq.	No.	Freq.
Complete Argument	9/22	40.9%	4/12	33.3%	5/10	50.0%
No Components	8/22	36.4%	5/12	41.7%	3/10	30.0%
Partial Argument	5/22	22.7%	3/12	25.0%	2/10	20.0%
Total Number Students	22	100%	12	100%	10	100%

Table 4: Post-Assessment Transition Frequencies

Pre-Assessment	Post-Assessment	Combined Courses		Semester Course		Summer Course	
		No.	Transition Frequency	No.	Transition Frequency	No.	Transition Frequency
NONE		N=8		N=5		N=3	
	No Growth	1/8	12.5%	0/5	0%	1/3	33.3%
	Partial Growth	6/8	75%	5/5	100%	1/3	33.3%
	None to b	2		2		0	
	None to ab	4		3		1	
	Complete	1/8	12.5%	0/5	0%	1/3	33.3%
	None to abcd	1		0		1	
PARTIAL		N=5		N=3		N=2	
	No Growth	1/5	20%	0/3	0%	1/2	50%
	d to d	1		0		1	
	Partial Growth	2/5	40%	1/3	33.3%	1/2	50%
	a to ad	1		0		1	
	d to ad	1		1		0	
	Complete	2/5	40%	2/3	66.7%	0/2	0%
	ab to abd	1		1		0	
	c to abc	1		1		0	

Table 4 provides the post-assessment transition descriptions and frequencies. For example, the 4th data row of Table 4 indicates 2 students in the combined courses exhibited a pre-to-post argument description transition of "No Components" on the pre-assessment to a post-assessment description with only the argument component "b" (transition labeled as "none to b"). In examining the transition frequencies for the combined courses in Table 4 we note the following: (1) 75% of students with no upper and lower bounds argument components on the pre-assessment provided a partial upper and lower bounds argument description on the post-assessment and 12.5% provided a complete argument description, and (2) 40.0% of students with a partial argument on the pre-assessment provided a complete upper and lower bounds argument description on the post-assessment. In the semester course, it is important to note that 2/3 of the students with a partial argument description on the pre-assessment transitioned to a complete argument description on the post-assessment. This is in contrast to the summer course, where one half of the students with a partial pre-assessment description exhibited no growth on the post-assessment and the other half exhibited only partial growth.

Conclusions and Discussion

The effectiveness of using video examples in online courses to stimulate the growth of teachers' ability to recognize and describe upper and lower bounds arguments of students is evidenced by the fact that 2/3 of the semester course students transitioned from a partial to a full upper and lower bound argument description on the post assessment, and 2/3 of the summer course students transitioned from a recognizing no components of the upper and lower bounds argument description to a partial or complete argument description. Some teachers recognized the yellow rod as an upper bound and the purple rod as a lower bound, but did not attend to the detail of the child's argument that there was no rod in between, so that the yellow rod was the smallest upper bound and the purple rod was the largest lower bound. Although there was some growth in teachers' recognition of components of children's arguments after studying the videos, there is still a need for improvement. The research suggests that a video-based approach for teacher education has the potential to be effective, but that a single-unit intervention may not be adequate for developing satisfactory adaptive expertise with regard to this particular form of reasoning. Future studies might include interventions that give greater attention to the variety of arguments, partial and complete, that children naturally develop in the process of problem solving so that there may be increased opportunities for teacher evaluations of the validity of the arguments posed. With regard to online courses, research also is needed to investigate the role of threaded discussion as a tool to develop adaptive expertise in recognition of children's emergent mathematical reasoning and what kinds of scaffolds may serve to stimulate group discussions that address important aspects of the process as can be observed through studying video data.

Endnote

[1] Research supported by the National Science Foundation grant DRL-0822204, directed by C. A. Maher with G. Agnew, C. E. Hmelo-Silver, and M. F. Palius. The views expressed in this paper are those of the authors and not necessarily those of the National Science Foundation.

[2] The repository for the project, Video Mosaic Collaborative, is accessible at the website: http://videomosaic.org/

References

Bransford, J. D., Derry, S. J., Berliner, D. C., & Hammerness, K. (2005). Theories of learning and their roles in teaching. In L. Darling-Hammond & J. Bransford (Eds.), *Preparing teachers for a changing world: What teachers should learn and be able to do*. San Francisco: Jossey-Bass.

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences, 2*(2), 141–178.

Cobb, P., Wood, T., & Yackel, E. (1990). Classrooms as learning environments for teachers and resesarchers. In R. Davis, C. Maher, & N. Noddings (Eds.), *Constructivist views on the teaching and learning of mathematics: Journal for Research in Mathematics Education, Monograph No. 4,* 125–146. Reston, VA: National Council of Teachers of Mathematics.

Fennema, E., Carpenter, T. P., Franke, M. L., Levi, L., Jacobs, V. R., & Empson, S. B. (1996). A longitudinal study of learning to use children's thinking in mathematics instruction. *The Journal for Research in Mathematics Education, 27*, 403–434.

Greeno, J. G., & Middle School Mathematics through Applications Group. (1998). The situativity of knowing, learning, and research. *American Psychologist, 53*, 5–26.

Maher, C. A., & Davis, R. B. (1995). Children's explorations leading to proof. In C. Hoyles and L. Healy (Eds.), *Justifying and proving in school mathematics* (pp. 87–105). London: Mathematical Sciences Group, Institute of Education, University of London.

Maher, C. A., Mueller, M., & Palius, M. F. (2010). Using an informal learning environment to promote mathematical reasoning among middle-school students. *Mathematics Bulletin, 49*, 64–72.

Maher, C. A., Mueller, M., & Yankelewitz, D. (2009). A comparison of fourth and sixth grade students' reasoning in solving strands of open-ended tasks. In M. Tzekaki, M. Kaldrimidou, & H. Sakonidis (Eds.), *Proceedings of the 33rd Conference of the International Group for the Psychology of Mathematics Education, 4*, 73–80, Thessaloniki, Greece.

Mueller, M., Yankelewitz, D., & Maher, C. (2009). Eliciting student reasoning through problem solving. In S. L. Swars, D. W. Stinson, & S. Lemons-Smith (Eds.), *Proceedings of the 31st Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education, 5*, 811–818. Atlanta: Georgia State University.

Palius, M. F., & Maher, C. A. (2011). Teacher education models for promoting mathematical thinking. In B. Ubuz (Ed.), *Proceedings of 35th Conference of the International Group for the Psychology of Mathematics Education 1*, 321–328. Ankara, Turkey.

Spiro, R. J., Feltovich, P. J., Jacobson, M. J., & Coulson, R. L. (1992). Cognitive flexibility, constructivism, and hypertext: Random access instruction for advanced knowledge acquisition in ill-structured domains. In T. M. Duffy & D. H. Jonassen (Eds.), *Constructivism and the technology of instruction: A conversation* (pp. 57–75). Hillsdale, NJ: Lawrence Erlbaum Associates.

Tirosh, D. (2000). Enhancing prospective teachers' knowledge of children's conceptions: The case of division of fractions. *The Journal for Research in Mathematics Education*, *31*, 5–25.
Yackel, E., & Hanna, G. (2003). Reasoning and proof. In J. Kilpatrick, G. W. Martin, and D. Schifter (Eds.), *A research companion to Principles and Standards for School Mathematics* (pp. 227–236). Reston, VA: National Council of Teachers of Mathematics.
Yankelewitz, D., Mueller, M., & Maher, C. (2010). A task that elicits reasoning: A dual analysis. *Journal of Mathematical Behavior*, *29*(2), 76–85.

Van Zoest, L. R., Lo, J.-J., & Kratky, J. L. (Eds.). (2012). *Proceedings of the 34th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education.* Kalamazoo, MI: Western Michigan University.