Teachers Can Learn to Attend to Students' Reasoning Using Videos as a Tool

Carolyn A. Maher
Marjory F. Palius
James A. Maher
Cindy E. Hmelo-Silver
Robert Sigley
Rutgers University

There is a need for research in STEM (science, technology, engineering, and mathematics) teacher education that addresses the challenge of building teachers' pedagogical skills in fostering the development of mathematical reasoning in students. The Common Core State Standards for Mathematics provide teachers with guidance on how to promote mathematical practices that emphasize reasoning and justification through problem solving and that encourage an exploration of viable strategies, through mathematical modeling and facilitating communication in the classroom, to critique mathematical arguments (National Governors Association, 2010). For many teachers, these kinds of mathematical practices may not be what they experienced as learners, and, therefore, it is not clear to them how to engage their students in ways that enact the new Standards of Mathematical Practices.

As we and others have shown, digital video can be an excellent

Carolyn A. Maher is a distinguished professor and director of the Robert B. Davis Institute for Learning, Marjory F. Palius is associate director of the Robert B. Davis Institute for Learning, James A. Mahler is a researcher, Cindy E. Hmelo-Silver is a professor, and Robert Sigley is an applications programmer, all with the Graduate School of Education at Rutgers University, New Brunswick, New Jersey. Their e-mail addresses are carolyn.maher@gse.rutgers.edu, marjory.palius@gse.rutgers.edu, jmaher3@comcast.net, cindy.hmelo-silver@gse.rutgers.edu. and robert.sigley@gse.rutgers.edu

resource for improving teachers' skills at attending to students' mathematical reasoning (Borko, Jacobs, Eiteljorg, & Pittman, 2008; Derry, Hmelo-Silver, Nagarajan, Chernobilsky, & Beitzel, 2006; Sherin & Han, 2004; Van Es, 2009; Zhang, Lundeberg, Koehler, & Eberhardt, 2011). Video can show the complexity of classroom practice and make student and teacher thinking visible (Brown, 1992; Miller & Zhou, 2007). Part of what makes video so compelling is the degree of complexity that it can convey. It also can provide opportunities for teacher professional development by encouraging teachers to focus on teaching and learning in ways that they would not be able during classroom instructional time and by providing virtual experiences that allow detailed studies of student thinking (Borko et al., 2008; Miller & Zhou, 2007; Palius & Maher, 2011). Several researchers have argued that video can provide a context for productive discussion and reasoning about teaching and learning (e.g., Borko et al., 2008; Zhang et al., 2011). Derry et al. (2006) argue that one of the reasons for the potentially powerful effects of video is that it provides opportunities to make connections between theoretical ideas and real-world practice. A range of pedagogical approaches for using video for learning includes lesson study, video clubs, problem-solving cycles, and problem-based learning (e.g., Borko et al., 2008; Maher, Landis, & Palius, 2010; Zhang et al., 2011).

Why, then, has there been relatively little change in instruction that fails to recognize the power of student reasoning? One reason may be the lack of awareness of the existence of videos of children engaged in sophisticated reasoning for use in teacher education. We begin to explore the question by first examining teacher knowledge about student reasoning. Opportunities to study children's doing and talking about mathematics and providing convincing arguments for their solutions to problems might not typically be accessible to pre-service teachers or to many in-service teachers whose approach to instruction misses opportunities to observe how children learn and do mathematics. Video, however, holds the promise of providing a window into alternative classroom settings in which communication, collaboration, and the sharing of ideas are the norm. The use of video clips of children thoughtfully engaged in doing mathematics, thus, offers a new lens through which to view student learning and brings forth a question that guides our work: Does teacher study of certain videos improve their ability to recognize the variety of forms of reasoning used by the children?

We have been conducting research in teacher education using the problem-solving tasks and videos from prior long-term research on the development of mathematical reasoning in students. This work is part of the Video Mosaic Collaborative, which also makes videos, tasks, and related resources available via the Internet (www.videomosaic.org). We report the results of quasi-experimental studies conducted over three years with pre- and in-service teachers. The underlying hypothesis was that a particular video collection can serve as a pedagogical tool for deepening teachers' awareness of how students' mathematical reasoning can emerge naturally through problem solving when appropriate conditions have been established in the learning environment (Maher, 2008). This hypothesis is premised on the notion that teachers' ability to recognize children's reasoning is likely essential for tackling the bigger challenge of subsequent change in teaching practice. Thus, we start to address that challenge by investigating our hypothesis about the pedagogical value of certain videos.

The Video Mosaic Collaborative (VMC)

A quarter-century of research on the development of mathematical ideas and reasoning has yielded a video collection that features students engaged in mathematical problem solving across multiple content strands in classroom and informal settings (Maher, 2009; Maher & Martino, 1996; Mueller & Maher, 2009). The videos are an outcome of research that followed the same students over time and that shows their making sense of problems and persevering in solving them. The students in the videos use appropriate tools and construct personally meaningful representations that support them in reasoning abstractly and quantitatively (Maher, Powell, & Uptegrove, 2010). The videos also illustrate how the researcher, in the role of classroom teacher, facilitated interactions among students who worked in small groups, as well as in whole-class discussions, in ways that supported students' articulation of mathematical arguments and consideration of whether those arguments were convincing as justification for solutions to problem tasks. Multiple cameras were used to capture the talk and inscriptions that children produced while working on cognitively challenging, yet accessible, tasks that allowed them to explore mathematical ideas before receiving formal instruction on those topics in their regular school curriculum.

Well-documented examples of students' mathematical reasoning, which were initially identified through research conducted by many scholars, have now been prepared by VMC and populate a searchable database, the Video Mosaic Collaborative. Of particular relevance to the research in teacher education reported here are the videos in the counting-combinatorics strand, which come from a seminal longitudinal study that followed the same group of students from early elementary grades through high school and beyond, with follow-up interviews. The tasks and videos have been used in a variety of intervention designs with pre-service and in-service teachers at the elementary, middle, and secondary levels. We report on the results of the various interventions, through which teachers first became engaged as learners, by working on the tasks, finding viable arguments, considering the mathematical structure, and then studying videos of students' working on the same tasks. The student reasoning from this strand and the collection of tasks have been carefully documented (Maher, Powell, & Uptegrove, 2010). The video collection, which includes additional content strands, can be found in the VMC database (Agnew, Mills, & Maher, 2010).

Theoretical Framework

We ground our work in a theoretical framework that assumes that people learn best when actively engaged in interpreting the world (Bransford, Derry, Berliner, Hamerness, & Darling-Hammond, 2005; Palincsar, 1998). Video can help to provide a bridge that connects prior learning to new knowledge. Video allows teachers to have a virtual entrée into the world of the classroom (Sherin & Han, 2004). Because video is "not live," it enables one to rewind, re-watch, review, and reflect. In this way, it provides teachers with opportunities for analytic thinking that live classroom observations or clinical interviews cannot. Notably, it can provide a shared focal representation for professional development (Borko, Koellner, Jacobs, & Seago, 2011). In addition, video offers learners the opportunity to discern what is important in a particular situation and can provide a basis for comparison with their own lived experiences. In the case of the research reported here, video provides opportunities for teachers to see that learners are capable of engaging in sophisticated reasoning as they develop and apply their new knowledge of mathematics learning to authentic contexts. The use of video in professional development encourages the understanding of learning "through careful observation of students and their work" (Bransford et al., 2005, p. 79).

The interventions reported here build on the assumption that observation and analysis of student mathematical behavior can reveal learners' developing knowledge and ability to reason. While we sought evidence that teachers can build knowledge about students' mathematical reasoning from studying videos in a facilitated learning context, we also point to the potential for transferring what is learned through virtual means into attending better to students' reasoning in live classroom environments.

The Intervention

The design of the interventions typically was based on one of two models (Palius & Maher, 2011), with variations made to accommodate specific learning goals for the population of participants in particular university courses, and the interventions shared common features. Participants in the experimental classes worked in groups to solve counting problems, watched VMC videos of children who were solving the same problems, and then discussed the variations among the teacher and student solutions. Another common feature of interventions was the focus on students' mathematical reasoning, particularly how students' representations and models became tools that they could use in trying to express convincing arguments for their solutions. Although all intervention enactments used the same basic design, the duration of an intervention varied, and course instructors were free to adapt the interventions to their specific circumstances; analysis of these adaptations is in progress. All experimental participants worked on the same core problem strand and were given a subset of the same VMC videos of students' reasoning to view. In contrast, the comparison classes neither worked on the problems nor analyzed student arguments from videos as part of their course curriculum.

Methodology

Participants

The participants were 177 pre- and in-service mathematics teachers. The experimental group had 127 participants, and the comparison group had 50. The distribution of participants is shown in Table 1. Experimental groups included K-5 pre-service teachers, K-8 in-service teachers, and pre-service secondary teachers, with corresponding comparison groups.

The K-5 pre-service teachers at a private university in New Jersey were in classes taught by the same instructor. These classes were selected at random to be designated as experimental or comparison. The

Table 1
Distribution of Study Participants

Participant Group	Experimental (n = 127)	Comparison (n = 50)
K-5 Pre-service Teachers	47	25
K-8 In-service Teachers	54	14
Pre-service Secondary Teachers	26	11

experimental and comparison K-8 in-service teachers were participants from over 20 New Jersey school districts. The pre-service secondary teachers were mathematics majors in a state university in New Jersey. The comparison groups were drawn from teacher populations similar to those of the experimental groups. Four mathematics teacher educators (MTEs) led the experimental classes; three worked with in-service teachers, two of whom also worked with pre-service secondary teachers; and a fourth worked with pre-service elementary teachers. Four MTEs led the comparison classes, with one instructing in-service teachers, one instructing pre-service secondary teachers, and two instructing pre-service elementary teachers.

Assessments

We were interested in studying the extent to which the study participants noticed children's mathematical reasoning from a specific video clip, which was used in all the interventions in the counting strand and administered as a pre-test and post-test to both the experimental and comparison groups of teachers. The video clip was excerpted from a small group interview with four children and the lead researcher. The interview was conducted after the children had worked with a partner on the task of building towers of a particular height (e.g., 4-tall, 5-tall) when selecting from two different colors of Unifix cubes. Children used different representations and strategies and had varying ways of justifying the solutions they found. Thus, the purpose of the interview was to give the children an opportunity to share their own ways of thinking and to hear about what others did to solve the problem. This video was useful for assessment because study participants had the opportunity to recognize different mathematical arguments, which they were asked to describe in detail in an open-ended response. Participants also were provided with a transcript of the video, and the assessment prompt indicated that they could refer to the transcript to provide specific details about the children's arguments. Note that this particular video was used for only for assessment purposes; it was not used by any MTEs in the design of an intervention for an experimental class.

Coding

Our research team developed a detailed rubric to code the open-ended responses to the video-based assessment on reasoning (see Appendix). Each mathematical argument was broken down into its constituent features. This allowed us to measure which argument features were identified on an individual's pre-test and post-test and to determine how, if at all, what the individual noticed and described might have changed over the course of the intervention. In the video, the children discussed the patterns that they could see from modeling, using the Unifix cubes, and they used those patterns to support their mathematical arguments for convincing one another that their solutions were valid. Two arguments took the form of reasoning by cases: Case I consisted of five different cases, and Case II, a more elegant argument, used four cases. The third argument used inductive reasoning. For the cases arguments, the scoring rubric enabled us to code for which specific cases were described by study participants and, thus, whether they partially or completely described each of the two arguments by cases. Similarly, for the inductive argument, the rubric enabled us to code separately for the presence of its two constituent features (i.e., establishing that, with two colors of cubes, there are two possibilities for a tower of height one and that each tower then has two possible choices for the color of a cube to be added on for a tower of height two, and so on).

To illustrate the coding, we provide examples of the written responses by Subject A to the pre-assessment and post-assessment. As can be seen below, the pre-assessment shows an incomplete description of the induction argument, as it focuses exclusively on presenting a description of the numerical patterns mentioned by the children in the video and does not specify either of its constituent features. There is no mention at all of a cases argument. By contrast, the post-test reveals a complete description of the inductive argument and a complete description of the Case I argument. It also shows a glimmer of recognition of the Case II argument, as seen in the observation, "The others wanted to change her pattern," yet, because what constituted that alternate case (i.e., exactly three towers with two blue cubes and one red cube) was not mentioned, it could not be coded as complete.

Example of Pre-assessment: Subject A

One argument was to start with one cube and to see how many towers could be made. Then two cubes were used, and an argument was made that you would have four. The child manipulated the cubes and decided that three cubes would be eight, and then she tried to use a pattern—four would be 12. She followed a pattern but didn't manipulate the cubes and was incorrect.

Another argument made was that you start with two blocks high and then (2 x 2) make four towers; three blocks high would be 2 x 2 x 2 = 8; four blocks high would be 2 x 2 x 2 x 2 = 16. This was a convincing argument.

By using the patterns, the students were able to check their solutions. They could see that they showed all the possibilities.

Example of Post-assessment: Subject A

Milin and Michelle were using inductive reasoning to answer the tower problem. Both of them reasoned that, each time you added a block to the towers, you needed to multiply by 2 because there were only two colors. "Towers of one would be two towers; towers of two would be four towers because you would take the towers of one and add either a red or blue . . . "

Stephanie proved her answers by using proof by cases. "She started with no blue, then one blue, then two blues stuck together, then two blues apart, then all blue." Stephanie liked to show a pattern but resisted when the others wanted to change her pattern. She had a specific way that she wanted to show the towers, but she was able to convince Jeff that she had all possible combinations.

Ultimately, Jeff was convinced by Milan and Michelle. He would clearly state the inductive reasoning they used. All children were able to answer additional tower questions without actually building the towers.

"I am convinced!"

Analysis

Video assessment data collected from intervention contexts and comparison groups were scored blindly as an aggregated data set. Each assessment was coded by two scorers who worked independently, and we achieved inter-rater reliability of 90% or greater. Assessment responses were scored by the researchers by whether or not the study participants provided a complete description for each of the three argument types (two different cases arguments and an induction argument). The coded data were then analyzed. For analysis purposes, a study participant was reported as exhibiting growth on the post-assessment if the participant provided a complete description of an argument type that was not included in the participant's pre-assessment.

Results

Comparability of Groups

Although we recognize that the different groups of teachers have different experiences and content backgrounds, we did not know their knowledge of children's reasoning, as held prior the study. Therefore, their pre-assessments were analyzed to determine whether experimental and comparison participants were comparable before the intervention. Table 2 presents the pre-assessment complete argument description rates for the various categories of study participants. Specifically, 22.8% of the 127 experimental participants and 24.0% of the 50 comparison group participants provided a complete argument description of at least one student solution to the counting task. Of the 127 experimental participants, the pre-assessment complete argument rates were 12.8% for the K-5 pre-service participants, 23.1% for the secondary pre-service participants, and 31.5% for the K-8 in-service teacher participants. The pre-assessment complete argument description rates were analyzed for the four study subgroups to determine whether experimental and comparison participants were comparable before the intervention. A contingency table analysis was performed to test the null hypothesis of no difference in pre-assessment complete argument description rates for the comparison and experimental groups. The data in Table 2 indicate that there were no significant differences among groups at the $p < 0.05$ level (χ^2 (3) = 4.97, p = 0.17).

Combining Comparison Group Data

Table 3 contains the post-assessment growth rates for the three subgroups of comparison teachers. The subgroup growth rates varied from 0.00 to 0.08, with an overall mean of 0.04. That is, of the 50 participants in the aggregate comparison group, only 4% exhibited growth on the post-assessment. An ANOVA found that there was no significant

Table 2
Pre-assessment Complete Argument Descriptions

	Complete Task Descriptions		
Pre-Assessment Participant Category	0	1 or 2	Total
Comparison	38 (76.00%)	12 (24.00%)	50
Experimental	98 (77.17%)	29 (22.83%)	127
K-5 Teachers, Pre-service	41 (87.23%)	6 (12.77%)	47
K-8 Teachers, In-service	37 (68.52%)	17 (31.48%)	54
Secondary Pre-service	20 (67.92%)	6 (23.08%)	26
Overall Total	136 (76.84%)	42 (23.16%)	177

Table 3
Mean Growth Estimates for the Categories of Comparison Participants

Comparison Subgroup	*n*	Mean Growth Estimate	Lower 95% Growth C.I.	Upper 95% Growth C.I.
K-5 Teachers, Pre-service	25	0.08	0.02	0.25
K-8 Teachers, In-service	14	0.00	0.00	0.22
Secondary Pre-service	11	0.00	0.00	0.26
Combined Comparison	50	0.04	0.01	0.13

difference among the comparison subgroups at the $p < 0.05$ level ($F(2, 47) = 1.02$, $p = 0.37$). Based on the comparability of the post-assessment growth rates for the three comparison subgroups of participants, the data from all three were combined in the subsequent analysis of the growth rate of the three experimental subgroups of participants.

Experimental versus Comparison Groups

The third column of Table 4 contains the post-assessment growth rates of the comparison (aggregate) and the three experimental subgroups. As seen in Table 4, 51.85% of the K-8 in-service experimental teachers, 38.46% of the secondary pre-service experimental teachers, and 17.02% of the K-5 pre-service experimental teachers exhibited growth on the post-assessment. This can be contrasted with 4% of the comparison teachers who exhibited growth. A logistic regression analysis was performed to test the hypothesis of no difference in growth rate of each subgroup of experimental teachers in contrast with that of the comparison teachers. The analysis indicated that the data provided evidence that the growth rate of each of the experimental subgroups exceeded that of the comparison group at a significance level that varied from 0.05 to less than 0.0001.

As indicated in Table 4, the results of the logistic regression analysis are as follows: (a) the K-8 in-service experimental participants have 25.85 times the odds of growth compared to the comparison participants; (b) the secondary pre-service experimental participants have 15 times the odds of growth compared to the comparison teachers; and (c) the pre-

Table 4
Growth Rate of the Elementary Pre-service Experimental, Secondary Pre-service Experimental and In-Service Experimental versus the Comparison Group

Study Group	n	Growth Rate	Odds Ratio[a]	Odds Ratio Lower 95% CI	Odds Ratio Upper 95% CI	$\chi^2(1)$	*p*
Comparison	50	0.0400					
Exp. K-8 Teachers, In-service	54	0.5185	25.85	7.01	168.13	17.78	<0.0001
Exp. Secondary Pre-service	26	0.3846	15.00	3.50	104.60	10.73	0.001
Exp. K-5 Teachers, Pre-service	47	0.1702	4.92	1.15	33.88	3.78	0.052

Note. a. Odds of growth of each experimental group divided by odds of growth of comparison group.

service experimental participants have 4.92 times the odds of growth compared to the comparison participants.

Table 5 provides a breakdown of the growth rates in the experimental group by type of argument. Overall, the K-8 in-service teachers achieved the highest growth rate (51.85%), followed by the pre-service secondary (38.46 %), and pre-service K-5 teachers (17.02 %). It should be noted that both the two cases and inductive argument types contributed to the growth rates. A complete argument description of Case Argument 1 was the leading contributor to growth among experimental participants. It is interesting to note that at least 70% of the overall growth rate is attributed to the Case Argument 1 for each of the experimental subgroups. It is even more interesting to note that, while the overall growth rate varied among the three experimental subgroups, nearly 40% of the growth rate can be attributed to multiple argument type descriptions for each of these three subgroups of participating teachers.

Conclusions

We are encouraged, based on this research, to find growth in both pre and in-service teachers and at both elementary and secondary levels, in terms of teachers' identifying students' reasoning from a video. This suggests that interventions using the VMC tasks and videos have the potential to help teachers to recognize student reasoning, an important goal in the learning and teaching of mathematics. Because the growth rates vary, it may be useful to explore what may account for some of the differences.

We first address the finding that the Case 1 Argument was the major contributor to the overall growth rate for all of the experimental subgroups. Each of those participants had an opportunity to work on the task of building towers of a specified height (i.e., 4-tall, 5-tall, … *n*-tall)

Table 5
Growth Rate in Experimental Groups for Each of the Three Arguments in the Assessment Video

Argument Type	K-8 In-service	Secondary Pre-service	K-5 Pre-service
Case Argument 1	38.9%	26.9%	14.9%
Case Argument 2	13.0%	19.2%	4.3%
Inductive Argument	24.1%	11.5%	4.3%
Multiple Arguments	20.4%	15.4%	6.4%
Overall	51.85%	38.46%	17.02%

when selecting from two colors. The task calls for providing a convincing argument that all possible towers have been found, which requires using some kind of pattern to organize the solution. The use of cases tends to arise naturally in the reorganization of data from patterns. The assessment video features a student, Stephanie, who argues strongly for her particular organization of cases, namely, the Case 1 Argument. Although other students in the video question Stephanie about her organization and suggest a more elegant organization, namely, the Case 2 Argument, this was not described by as many participants. However, the secondary pre-service teachers provided this observation in their response most frequently, followed by the in-service teachers. The explanation of an inductive argument offered by students in the video was not articulated as clearly as was the Case 1 Argument and, thus, required more attentive listening to make sense of what the students were trying to say, as they elaborated on the pattern that they noticed and why it worked, to explain the doubling of the number of towers as their height increased by one. We suspect that the in-service teachers' greater experience in listening to children express their mathematical ideas contributed to that subgroup of experimental participants' having a higher growth rate.

It is notable that, consistently, the in-service teachers performed better. We argue that this occurred, perhaps, due to the particular model used in the in-service intervention (Maher, Landis, & Palius, 2010); this model enabled the teachers to carry out the tasks with their own students and to study the differences in arguments posed by their students. This direct feedback was not available to the pre-service teachers, who, as novices, have little or no opportunity to work directly with students. The better performance by in-service teachers leads us to conjecture that the demonstrated ability to understand and explain student reasoning via virtual means could lead to an increased capacity to do so in the classroom. This is a worthwhile area for future research.

The instances of the higher growth rate for the secondary pre-service teachers might be explained, in part, by their stronger mathematical backgrounds. For example, they out-performed the in-service teachers in description of the Case 2 Argument due to their paying greater attention to the inelegance of the Case 1 Argument. Not surprisingly, their growth rates were higher for all arguments than for those of the elementary pre-service teachers. Yet, it is interesting to note that a background as a mathematics major was not sufficient to enable the teacher to recognize children's emergent reasoning. Even secondary pre-service teachers had room to grow from pre- to post-assessment and improved in recognizing more detail in the children's arguments, even though their intervention was of the shortest duration.

More work needs to be done at the elementary level. Our findings show that challenges have been identified (e.g., Ball, Hill, & Bass, 2005) and that other researchers have explored ameliorating these challenges through the use of video case studies and focused discussions (Borko et al., 2008; Llinares & Valls, 2009). With regard to fostering pre-service elementary teachers' recognition of children's mathematical reasoning, it may turn out that working with students as part of the intervention could have a substantial payoff. Future research could explore value of models that provide opportunities to move from theory and knowledge into direct practice. Perhaps such research could be accomplished through small teaching experiments conducted by pre-service teachers with students, using the tasks and videos as tools to design interventions and study the developing reasoning of the students in both informal and formal settings. Further work also could follow some of these pre-service teachers into their student teaching practicum to look for sustained effects that are visible only on a long-term basis. Another worthwhile line of research is to investigate the varied details of an intervention (e.g., number of tasks, amounts of time spent on face-to-face/online activities, individual/class viewing) by instructors as a means to shed more light on what underlies the differences in growth rate among treatment groups. This work is currently in progress, yet some implications of the research presented here are already apparent.

The analysis of the assessment data from interventions in the counting strand, which showed growth rates that range from 17% to 52%, as contrasted with 4% for the combined comparison groups, provide evidence that these interventions can be effective in helping teachers to learn to attend to students' mathematical reasoning. Further, despite the small sample size, we found differences between experimental and comparison groups. The positive results from our design research studies contribute to the literature on mathematics teacher education and professional development. That the strongest growth was found among the in-service teachers points to the value of an intervention model that includes classroom-based task implementation in addition to teachers' problem solving and studying videos of students engaged in solving those same tasks. Although our model has activities also used in other approaches to professional development, such as problem-solving cycles (Koellner et al., 2007), the sequence of activities and the video resources used are different.

The VMC videos feature a researcher in the role of a classroom teacher who is an expert in facilitating classroom discourse among students engaged in cognitively challenging tasks (Palius & Maher, 2011). Studying these videos, after working on the tasks themselves but before doing classroom implementation, may be particularly useful for teachers who are learning how to align their practices with the Common Core State Standards related to mathematical content and practices. The VMC videos show students engaged in making sense of problems as well as how the teacher fosters the norm of sense making in the classroom. These videos also show how students attend to structure and make use of it in their reasoning as well as how the teacher engages students in mathematical discussions to construct arguments and critique the reasoning of others. Perhaps of greatest importance to mathematics teacher educators is that the resources of the VMC, which include statements of problem-solving tasks as well as videos, are openly accessible (www.videomosaic.org) and can be utilized broadly in courses and professional development programs. Looking more broadly at STEM education, the resources, models, and tools that have emerged from the VMC research, along with the promising findings from their applications to teacher education, give further support to the medium of video for use by teacher educators to deepen teachers' understanding of standards-based practices.

Note

The Video Mosaic Collaborative is a research and development project sponsored by the National Science Foundation (Award DRL-0822204). We gratefully acknowledge the support of the National Science Foundation and note that the views expressed in this paper are those of the authors and not necessarily those of the NSF.

References

Agnew, G., Mills, C. M., & Maher, C. A. (2010). VMCAnalytic: Developing a collaborative video analysis tool for education faculty and practicing educators. In R. H. Sprague, Jr. (Ed)., *Proceedings of the 43rd Annual Hawaii International Conference on System Sciences (HICCS-43): Abstracts and CD-ROM of full papers*. Los Alamitos, CA: IEEE Computer Society, Conference Publishing Services.

Ball, D. L., Hill, H. C., & Bass, H. (2005). Knowing mathematics for teaching: Who knows mathematics well enough to teach third grade, and how can we decide? *American Educator, 29*(1), 14-17, 20-22, 43-46.

Borko, H., Jacobs, J., Eiteljorg, E., & Pittman, M. (2008). Video as a tool for fostering productive discussions in mathematics professional development. *Teaching and Teacher Education, 24*, 417-436.

Borko, H., Koellner, K., Jacobs, J., & Seago, N. (2011). Using video representations of teaching in practice-based professional development programs. *ZDM, 43*(1), 175-187.

Bransford, J. D., Derry, S. J., Berliner, D., Hamerness, K., & Darling-Hammond, L. (2005). Theories of learning and their roles in teaching. In L. Darling-Hammond & J. D. Bransford (Eds.), *Preparing teachers for a changing world* (pp. 40-87). San Francisco, CA: Jossey-Bass.

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

Derry, S. J., Hmelo-Silver, C. E., Nagarajan, A., Chernobilsky, E., & Beitzel, B. (2006). Cognitive transfer revisited: Can we exploit new media to solve old problems on a large scale. *Journal of Educational Computing Research, 35*, 145-162.

Koellner, K., Jacobs, J., Borko, H., Schneider, C., Pittman, M. E., Eiteljorg, Bunning, K., Frykholm, J. (2007). The problem-solving cycle: A model to support the development of teachers' professional knowledge. *Mathematical Thinking and Learning 9*(3), 273-303.

Llinares, S., & Valls, J. (2009). The building of pre-service primary teachers' knowledge on mathematics teaching: interaction and online video case studies. *Instructional Science 37*, 247-271.

Maher, C. A. (2008). Video recordings as pedagogical tools in mathematics teacher education. In D. Tirosh & T. Wood (Eds.), *International handbook of mathematics teacher education: Vol. 2. Tools and processes in mathematics teacher education* (pp. 65-83). Rotterdam, The Netherlands: Sense.

Maher, C. A. (2009). Children's reasoning: Discovering the idea of mathematical proof. In M. Blanton, D. Stylianou, & E. Knuth (Eds.), *Teaching and learning proof across the grades: A K-16 perspective* (pp. 120-132). Mahwah, NJ: Taylor & Francis Group.

Maher, C. A., Landis, J. H., & Palius, M. F. (2010). Teachers attending to students' reasoning: Using videos as tools. *Journal of Mathematics Education, 3*(2), 1-24.

Maher, C. A., & Martino, A. M. (1996). The development of the idea of mathematical proof: A 5-year case study. *Journal for Research in Mathematics Education, 27*(2), 194-214.

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., Powell, A. B., & Uptegrove, E. (Eds.), (2010). *Combinatorics and reasoning: Representing, justifying and building isomorphisms*. New York: Springer.

Miller, K., & Zhou, X. (2007). Learning from classroom video: What makes it compelling and what makes it hard. In R. Goldman, R. Pea, B. Barron, & S. J. Derry (Eds.), *Video research in the learning sciences* (pp. 321-334). Mahwah NJ: Lawrence Erlbaum Associates.

Mueller, M., & Maher, C. A. (2009). Learning to reason in an informal math after-school program. *Mathematics Education Research Journal, 21*(3), 7-35.

National Governors Association Center for Best Practices and Council of Chief State School Officers. (2010). *Common Core State Standards Initiative, Mission statement*. Retrieved from http://www.corestandards.org/

Palincsar, A. S. (1998). Social constructivist perspectives on teaching and learning. *Annual Review of Psychology, 45*, 345-375.

Palius, M. F., & Maher, C. A. (2011). Teacher education models for promoting mathematical thinking. In Ubuz, B. (Ed.), *Proceedings of 35th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 321-328). Ankara, Turkey: PME.

Sherin, M. G., & Han, S. Y. (2004). Teacher learning in the context of a video club. *Teaching and Teacher Education, 20*, 163-183.

Van Es, E. (2009). Participants' roles in the context of a video club. *Journal of the Learning Sciences, 18*, 100-137.

Zhang, M., Lundeberg, M., Koehler, M. J., & Eberhardt, J. (2011). Understanding affordances and challenges of three types of video for teacher professional development. *Teaching and Teacher Education, 27*, 454-462.

Appendix
VMC Studies: Counting Video Assessment Scoring Rubric

Scoring Holistically

Study participants watch a video clip from the "Gang of Four" interview with the researcher and four 4th graders: Milin, Michelle, Jeff, and Stephanie. In an open-ended format, participants respond to a prompt that asks them to describe as completely as they can: (1) each example of reasoning that a child puts forth; (2) whether or not the reasoning forms a valid argument; (3) whether or not the argument is convincing; and (4) why or why not they are convinced. They are asked to provide evidence from the interview to support any claims that they make, and they are provided with copy of transcript for the video clip.

Scoring of an assessment begins by the researchers' reading the participant's response in its entirety to get a sense of its scope. Then it is reviewed more carefully to look for written evidence that supports scoring of particular rubric items. Care is taken due to the participant's being free to express his or her response in any desired organization within the open-ended response format. The entire response is thus considered, as a participant may respond to one part of the assessment instructions in detail and not repeat this detail in response to the other parts. The scoring focus is on mathematical reasoning, with less importance on the language used to express that reasoning. For instance, a sophisticated response may present the name of an argument type and a discussion of it only in general form; other responses may use very informal language. What someone says in his or her response matters more than how it is expressed.

Argument Forms and Constituent Features

Cases Argument 1: Stephanie's cases argument for towers three cubes high that are selected from two colors (blue and red) results in a set of eight unique towers. A complete argument includes each of the following cases. Note that written responses by study participants may well be fragmentary and use much less precise language than the following.

- All blue cubes or no red cubes, resulting in only one tower.

• One blue cube and two red cubes, resulting in three unique (different) towers.

• Two blue cubes stuck together and one red cube, resulting in two unique towers.

• No blue cubes or all red cubes, resulting in one tower.

• Two blue "stuck apart" or separated by one red cube, resulting in one tower.

Cases Argument 2: An alternate cases argument for towers three cubes high that are selected from two colors (blue and red) proposed by several of the children. Several of the cases overlap completely with the ones articulated by Stephanie. Participants may describe the organization of the third case as better (e.g., preferred, more elegant) than the way that Stephanie organized her cases, which bifurcated it into the third and fifth cases in the Cases Argument 1, above.

• All blue cubes or no red cubes, resulting in only one tower.

• One blue cube and two red cubes, resulting in three unique (different) towers.

• Two blue cubes and one red cube, resulting in three unique (different) towers.

• No blue cubes or all red cubes, resulting in one tower.

Inductive Argument: This argument may be expressed with reference to towers of a specific height, as in the two features below. It also may be expressed in general form.

• When building towers that are selected from two colors, there are exactly two unique towers of height one. With a single position in the tower, the one cube can be (say) either red or blue.

• Two unique towers of height one can be used to generate all possible towers of height two. For each tower one cube in height, two different towers can be built from it. Starting with (say) a red cube in the first position, either a red cube or a blue cube can be placed in the second position. Similarly, starting with a blue cube in the first position, either a red cube or a blue cube can be placed in the second position. The resulting four unique towers of height two is double the amount, two, that there are of towers of height one. (And so on for *n*-tall.)