Using Guided Experiences with Video to Help Teachers Interpret Children's Thinking with Appropriate Intellectual Modesty

Michael D. Preston
Teachers College, Columbia University
Columbia Center for New Media Teaching and Learning
mdpreston@columbia.edu

Paper presented at symposium
"Video Analysis as a Method for Developing Preservice Teachers' Beliefs About Teaching and Their Understanding of Children, Pedagogy, and Assessment"
2010 AERA Annual Meeting, Denver, CO
May 4, 2010

Many things guide good teaching: a robust understanding of content, a mastery of pedagogical techniques, an understanding of how children think and learn. Schools of education attempt to prepare teachers in each of these areas. Nevertheless, learning how to teach effectively often ends up happening on the job, when the teacher is dealing with real children in a complex classroom setting. In this new context, some of the ideas and skills teachers learn during their years of preparation may transfer well, such as how to design a lesson or how to teach a particular topic, but others are more elusive. This paper addresses one particularly challenging need: how to help teachers better understand what children know and can do, and to motivate them to learn more about their children as part of an ongoing process of formative assessment to guide teaching.

Learning to observe children

Using videotapes of children in classrooms to help prepare teachers is hardly a new concept (e.g., Fuller & Manning, 1973). But video alone is not sufficient. If the goal is to help teachers develop an "enlightened eye" (Eisner, 1998) for observing children—to discriminate what matters from what does not—it is necessary to "educate their perception" by first giving them the tools to discern and differentiate before introducing them to new content or new experiences (Schwartz & Bransford, 1998), confirming Piaget's (1976) notion that "…if they are not on the look out for anything … they will never find anything." Schwartz and Hartman (2007) suggest a useful approach they call "designed video," or the careful selection of video content aligned with a method of engaging students to meet specific learning goals, which might range from learning to perform a task, to merely attending to or understanding something, each of which puts different demands on the teacher and the learner.

The use of video for teacher education demonstrates the full spectrum of relatively "un-designed" to very designed video. Video has long been used to help teachers observe, assess and confront their own behaviors, but not necessarily to apply the same level of focus to children. Since the 1960s, researchers have utilized video to help teachers review and improve their own teaching practice (Fuller & Manning, 1973). The majority of studies on video-based methods of teacher development concentrate on lesson analysis, including an extensive body of work in mathematics education. These studies range in focus from what teachers attend to in the videos (Star & Strickland, 2008), to teachers' development of observation and reasoning skills (Santagata, Zannoni, & Stigler, 2007), and cross-cultural comparisons of teacher "noticing" about the features of videotaped lessons (Miller & Zhou, 2007).

The concept of "noticing" signifies a shift in observational focus from teacher behavior to child behavior (Sherin & van Es, 2005; van Es, 2008; van Es & Sherin, 2002). Following the logic that observation should include both perceptual and judgment components, efforts designed specifically to improve teacher noticing emphasize both improving (1) what teachers notice—from a focus on teacher action to a focus on children's conceptions—and (2) how teachers reason—from simply reporting what they see to more cognitively demanding but formatively useful activities such as synthesizing, generalizing, and interpreting (Sherin & Han, 2004).

The work on teacher noticing also converges with mathematics reform literature (NCTM, 2000; NRC, 2001), specifically, teacher investigation of and building upon children's thinking. Indeed, both tend to focus the interrelatedness of skills such as "attending to children's strategies, interpreting children's understandings, and deciding how to respond on the basis of children's understandings" (Jacobs, Lamb, & Philipp, 2009). Putting children at the center of the process should similarly change the emphasis of video-based instruction for teachers.

Learning to think critically about evidence

Having gathered some evidence, the next step is to adopt a critical approach and decide on what is meaningful. The evidence must be weighed; for example, does the evidence gathered truly reflect something about the child's understanding about what has been taught, or is it a procedural bug in the child's method of solving the problem? Or was the child simply distracted? Like Socratic questioning, this type of thinking requires teachers to make assumptions, distinguish between relevant and irrelevant information, and explain points.

"Critical thinking" is a term with many definitions, ranging from skills based in logical inquiry and reasoning (Glaser, 1985) to dispositions and abilities that contribute to better thinking (Ennis, 1987). Critical thinking involves specific activities, including recognizing one's assumptions and values, evaluating arguments and evidence, drawing inferences, and altering judgments when justified (Glaser, 1985). It can also be dialogical or dialectical (Felton & Kuhn, 2001), in the true Socratic sense, requiring refinement and revision to one's thinking in response to challenges from others.

There are explicit methods for teaching the skills of critical thinking, including such activities as interpreting texts, analyzing existing arguments, identifying the assumptions on which arguments are based, and even carefully evaluating one's own thinking (Ennis, 1987). However, a number of studies identify instances where these skills break down (Halpern, 1998; Kuhn, 1999). For example, students fail to apply the analytic and evaluative skills they have learned in a specific context to other contexts, or to their own thinking in general (Paul, 1995). Even when the skills of argument are explicitly taught, there is a good chance that teachers will not use them in everyday contexts. To a certain extent, whether an individual thinks critically about information in general depends on the development of personal epistemological beliefs.

An alternative to explicit training in the skills associated with critical thinking is to provide a context in which students can engage in regular practice, a strategy based on the simple notion that one can improve one's thinking simply by doing more of it (Dewey, 1910; Kuhn, Shaw, & Felton, 1997). Practice makes better, and not just any practice, but "deliberate" practice done with full concentration and aimed at generating improvement (Ericsson & Charness, 1994). Learning activities can be designed to improve performance in specific skills, they can be somewhat artificial and graduated, and they can include guidance and timely feedback. Kuhn suggests that a cognitively rich environment that sets up the conditions for engaging in argument can aid in the development of relevant skills (Kuhn, 2001; Kuhn, et al., 1997). Her research focuses on the impact of argument in social learning context, in collaboration with and/or in opposition to peers of roughly equal competence.

Another component to learning to think critically is developing an appropriate disposition toward evidence and a willingness to revise one's thinking. Otherwise, information as seemingly straightforward as a child's performance on a standardized assessment might be taken at face value, without further questioning that might reveal a deeper truth, such as actual understanding that was obscured by a simple flaw in the child's problem-solving method. A teacher oriented to think critically wants to know more, considers alternatives, and delays judgment when further investigation seems warranted. Two valuable conclusions might be drawn from the many taxonomies of critical thinking values, such as Edman (2008): first, to be a critical thinker means to have a tolerance for imperfect understanding, and second, good thinking requires a commitment to constant learning. Progress toward better understanding comes only from a continuous effort to acquire new information, even when it contradicts what one believes they already know or reveals even greater deficits of knowledge. Because belief preservation is a natural tendency, good critical thinkers compensate by putting extra effort into searching for and attending to evidence that contradicts their beliefs, and thereby cultivate a willingness to change their minds when the evidence starts mounting against those beliefs (van Gelder, 2005).

Close viewing study

A study was developed to test whether prospective teachers could be taught to use more interpretive language with respect to evidence, and whether they would also demonstrate the signs of "intellectual modesty" in their writing. Previous work had suggested that students benefit from a classroom-based method of guided interactions with video coupled with private study using a web environment that provides frequent, controlled opportunities to observe and interpret children's behavior in order to substantiate theories about how children think and should be taught (Ginsburg, Cami, & Preston, 2009). The study was designed to capture the salient features of this pedagogical method in a closed online space, where participants would undertake a series of lessons in early mathematics with video delivered in either a "guided" or "unguided" method described further below.

Sixty participants were recruited for the study. The participants were undergraduate students preparing to become early childhood or elementary teachers who had not yet taken a course in mathematics methods or begun student teaching, in order to ensure consistent levels of familiarity with the subject matter and teaching experience. The participants—57 female and 3 male—were recruited from 10 colleges and universities around the country, and participants from each institution were distributed evenly among the three conditions. They were informed that they were participating in the testing of new experimental lessons on early childhood mathematics, and would be compensated for their participation to ensure that they completed all seven sessions (participants in the control group completed only the pre- and post-tests and were therefore paid a commensurable amount for their time).

Participants were distributed randomly among the three conditions, N = 20 per group. The three conditions were designed as follows:

1. Guided video group: These participants experienced a "guided lesson" format, meaning that video segments were divided into short, meaningful segments, and each segment was followed by a question of one of the following types:
   - Describe what you observed; interpret the child's behavior;
   - Offer a hypothesis for what the child knows;
   - Predict what will happen next;
   - Provide further questions to ask or tasks to try;
   - Suggest strategies for teaching.

   Once participants submitted an answer, they were not permitted to change it. After answering each question, they received a list of sample answers to the question framed as "successful student responses," which were intended to provide them with a range of possible interpretations, and to model appropriately "modest" language. These responses were intended to give participants an opportunity to self-assess by benchmarking their responses against those of their peers. At the conclusion of each lesson, participants responded to the summative questions about what the child knew and what next steps they would take as the child's teacher. (A description the lesson design is presented below.)
2. Unguided video group: Participants in the unguided video group watched the same assigned video segments in their entirety, unsegmented and without interpretive questions to prompt them. Rather, before watching the video they received a list of potentially significant moments for which they should look carefully, but did not receive feedback after submitting responses. After watching the video, participants responded to the same summative questions as the guided group.
3. Control group: Participants in the control group completed only the pre- and post-tests, with a three-week delay in between sessions.

Participants were not told of the existence of other groups or the condition to which they were assigned.

Lessons in early childhood mathematics education

Five lessons related to the topic of "number" were developed for the experiment. Number was selected because it is the first of the NCTM recommended content areas

(NCTM, 2000) and generally the most readily accessible for teachers and students. The selected sub-topics within number—counting, enumeration, addition, subtraction, and equivalence—follow a developmental sequence in terms of mathematical complexity and the age at which children learn them, but the participants began the sequence from a randomized first lesson in an attempt to control for an order effect. Each lesson began with a reading of about two pages, derived from *Children's Arithmetic* (Ginsburg, 1989), to summarize the topic and help participants develop a sense of what to look for in the video. The videos selected for each lesson were about two minutes in length, and featured a child engaged in a mathematical activity or solving a problem related to the topic and discussing his or her thinking with an adult interviewer. These videos were carefully selected both for their relevance to the topic and for their complexity, in order to provide participants with an opportunity to interpret what they observed, and for the potential to have their initial assessments proved incorrect.

For the guided lesson condition, the videos were segmented into short clips of approximately 10 seconds in length, with exact length dependent on the content of the clip. An interpretive question followed. For example, a segment might be cut when the child in the video stated the answer to a question posed by the researcher, and then participants were asked to theorize about how the child arrived at the solution. Each lesson in the guided format included an average of 15 questions constructed in this way.

At the conclusion of each lesson, participants in both the guided and unguided groups responded to two summative questions about the video: an assessment of what the child knew about the topic, and a plan for what the participant would do next as a teacher. The lessons were designed to take 45-60 minutes total, including reading the text, watching the video, and responding to the questions. Participants in the guided and unguided conditions were required to complete two lessons per week, so their participation in the study lasted three to four weeks.

Participants in all three groups completed the pre- and post-tests, which used videos that were similar to those in the lessons but were longer (about six minutes) and designed as a transfer test by focusing on a different early mathematics topic, namely, the understanding of pattern. The pre- and post-tests included no introductory reading, and the video was unsegmented. At the conclusion of the video, participants were prompted to respond to the summative questions about what the child knew and what should be done next to teach the child. The tests employed two roughly equivalent videos on pattern, which were counter-balanced to control for an order effect. The pre- and post-tests were designed to take approximately 30 minutes.

Coding scheme

The study focused on data collected in the first summative question in the pre- and post-tests and at the conclusion of the five lessons. The data were coded according to the following scheme:

(1) CLAIMS: A claim is a generalization, a statement of belief, or an assertion. It can also be a prediction. A claim tends to introduce a new idea, e.g., “Children can count mentally or use a variety of other strategies.” A claim tends to be broad and to require substantiation, e.g., “The boys demonstrate a strong understanding of spatial relations.”

(2) EVIDENCE: Evidence is a reference to or description of observable events, usually positioned after a claim. In this study, evidence could appear as text (verbal description) or video inserted within the essay. Evidence contains observable events that two people can more or less agree upon objectively, e.g., “Armando tries to add another block to connect the two structures, but it doesn’t reach.” Evidence can include both verbal and nonverbal behaviors, e.g., “Armando begins looking around for a certain block and says ‘circle thing’ to describe it.” There is an interpretive component in the naming and placement of a clip, but any deliberate attempt to use evidence to support a claim constitutes an interpretation (see below). Evidence refers exclusively to the naming or identifying of observable behaviors.

(3) INTERPRETATIONS: An interpretation offers a plausible explanation for what is happening in the cited evidence, e.g., “Armando’s use of the phrase ‘circle thing’ demonstrates that he knows some shapes and can identify this aspect of the cylinder, even if he doesn’t have the proper word for it.” An interpretation explains how the evidence supports (or contradicts) a claim, e.g., “Gabriella appears to know what colors can be used in her blue-green pattern, but when she chooses yellow, it shows that she may be more focused on the colors than on the rules of patterns.” Interpretive keywords include “shows” and “demonstrates,” i.e., words the author uses to explain the evidence or to comment on something that is otherwise observational.

(4) MODESTY: Intellectual modesty is signified by “modals,” or expressions that are “used to qualify the truth of a judgment” (Garson, 2009). Specifically, this study was concerned with the language of epistemic modality, or words used to evaluate knowledge or a belief, or express confidence (or a lack of confidence) in the knowledge or belief upon which a proposition is based (Loos, Anderson, Day, Jordan, & Wingate, 2003). In speech, modality can be represented grammatically as a verb (e.g., may, might) or adverb (perhaps, possibly). For the purposes of coding, modest language includes explicit statements in which the author assesses the relative certainty of a specific interpretation, e.g., conditional words like “might” and “could” (anticipating other possible interpretations), perception words like “appears” and “seems” (limiting certainty), temporal words like “at this point” and “before” (acknowledging interpretations can change with new evidence), and metacognitive words like “we realize” and “leads one to believe” (inserting the author’s thinking into the essay). Modesty can also identify missing evidence, e.g., “Because the interviewer changed tasks, we did not see whether Gabriella could continue the pattern on her own.”

Results of the study

“Interpretation” and “modesty” in the guided video group were the only variables that demonstrated significant differences between the pre- and post-tests. There were no significant differences for “claims” and “evidence” in the guided video group, and no significant differences from pre-test to post-test for any variable in either the unguided video or control groups. These results suggest that the guided video condition did improve participants' interpretive skill and intellectual modesty, while the unguided video condition, as well as the control condition, did not.

Focusing exclusively on interpretations and modesty, a 3x2 analysis of variance was used to compare group means among all three conditions, to demonstrate the effect of the guided video condition. To demonstrate that these improvements in the guided group were relatively consistent among participants (and lack of improvement consistent among participants in the other two groups), Fisher's exact test was employed to compare scores that increased with scores that stayed the same or decreased. Fisher's exact test was used rather than a chi-square test because the expected value of some cells in the contingency table was below 5.

Comparing interpretations across the three groups in the study, there was a significant interaction ($p < .01$) between groups and repeated measures, as shown in Figure 1. The difference within groups was significant only for the guided video group ($p < .01$). The unguided video group showed a non-significant decrease ($p = .11$).

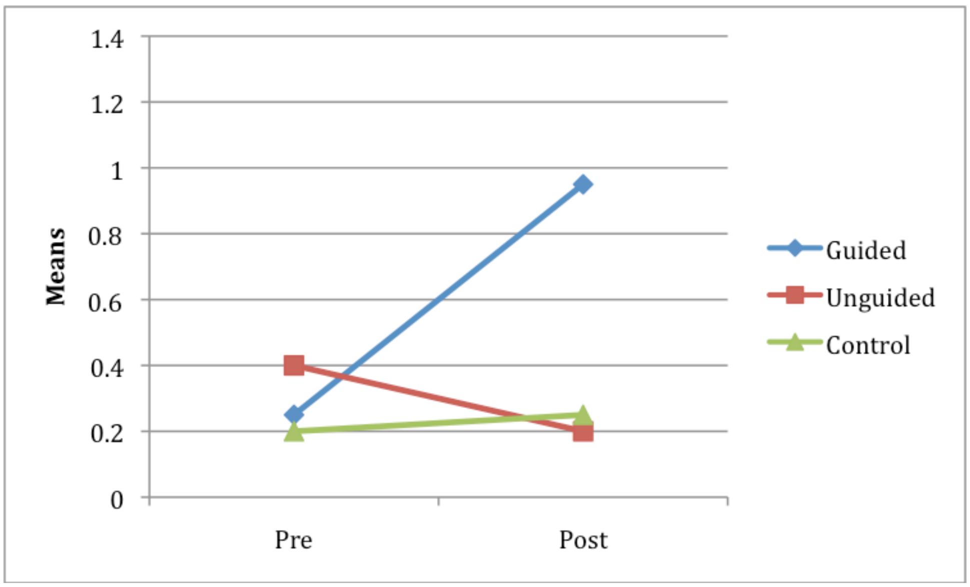

Figure 1: Pre/post comparison of interpretations for guided video, unguided video, and control groups.

Fisher's exact test was employed to show differences in scores within each group. Because the test requires a 2x2 plot, it was run for all three combinations of groups. The values in each cell represent counts of scores that either increased from pre- to post-test or did not increase. Scores that stayed the same or decreased were combined because the number of scores that decreased was generally very small. For interpretations, there were significant differences between the guided video group and the unguided video group and the control group, both at $p < .01$. The difference between the unguided video group and the control group was not significant.

These tests illustrate the difference in improvement between participants in the guided video group and those in the other two groups. More than half of the participants in the guided group increased their use of interpretations between the pre- and post-test, with 12 increased scores, one decreased score, and seven instances of no change. Comparatively, a much smaller number of participants in the unguided and control groups increased their use of interpretations. In the unguided group, there was one increased score, three decreased scores, and 16 instances of no change. In the control group, there were three increased scores, three decreased scores, and 14 instances of no change.

Comparing modesty across the three groups, there was a significant interaction ($p < .01$) between groups and repeated measures, as shown in Figure 2. The difference within groups was significant only for the guided video group ($p < .01$).

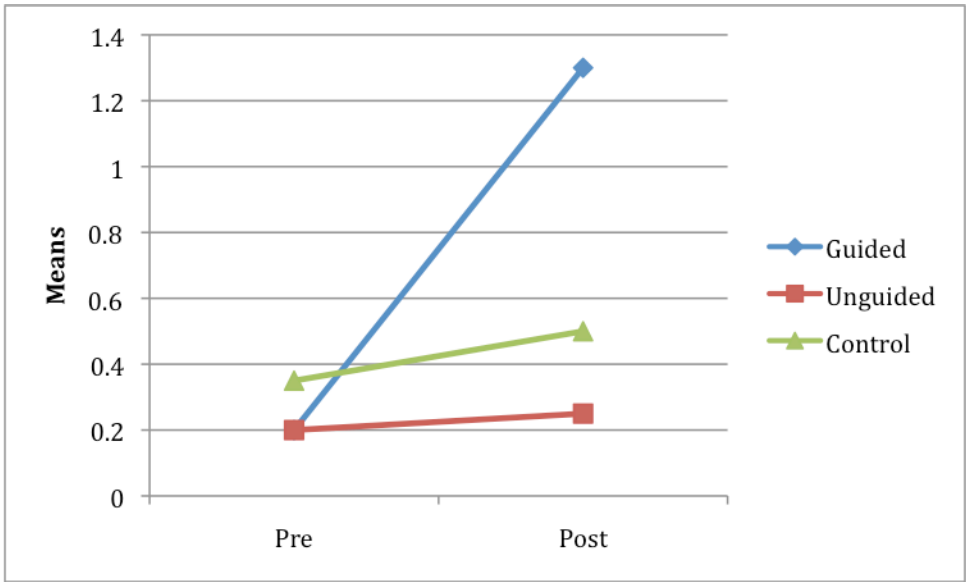

Figure 2: Pre/post comparison of modesty for guided video, unguided video, and control groups.

Comparing participants' scores using Fisher's exact test, there were significant differences between the guided video group and the unguided video group and the control group, both at $p < .01$. The difference between the unguided video group and the control group was not significant. As above, these tests illustrate a significant difference in improvement by participants in the guided video group when compared to participants in the other two groups. More than half of the participants in the guided group increased their use of modesty between the pre- and post-test, with no decreased scores, and nine instances of no change. Few participants in the unguided and control groups increased their use of modesty. In the unguided group, there were two increased scores, one decreased score, and 17 instances of no change. In the control group, there were five increased scores, four decreased scores, and 11 instances of no change.

These results show that the guided video group made significant gains in interpretation and modesty, and these gains were significantly different from the unguided video and control groups. Because the pre- and post-tests were the only consistent task for all participants in the study, they represent the most valid means of comparison between groups.

Data from the five lessons completed by the guided video and unguided video groups were also analyzed using the same coding scheme to determine whether there was any general improvement over the course of the study, and whether any specific lessons were more difficult than others. There was no improvement in performance detected in either group, and performance across the various lesson types was relatively consistent.

Discussion

The study suggests that the skills and dispositions of good thinking can be developed with regular practice in the guided viewing of videos. There are two key implications: first, that the educational impact of video changes substantively when the content is presented in such a way that it engages students in close viewing and active analysis, and second, that it is possible to change students' perceptions of certainty about their own knowledge by placing them in a situation in which evidence is complex, challenging, and changing, and assumptions must therefore be reconsidered and interpretations revised.

The first key finding is that the guided method of video analysis—segmented video, questions about each clip, and "feedback" in the form of sample responses to the questions—proved effective in increasing the likelihood that participants would produce reasonable interpretations of the evidence placed before them. The second key finding is that the guided method increased the likelihood that participants would employ language signifying "intellectual modesty," indicating a lack of certainty or a need for more information to support a claim. Taken together, these findings suggest that it is possible to affect both the skills and dispositions of critical thinking, specifically in the context of formative assessment techniques, including observation and interpretation of children's behavior, and willingness to repeat the process in order to refine one's understanding of a child's thinking.

Interpretation is defined as a reasonable explanation for what has been identified as evidence that supports a given claim. In the study, there was a significant increase in the use of interpretation by participants in the guided condition, and no significant change for participants in either the unguided or control conditions. Statistically significant differences were found only in the comparison of the pre- and post-tests and not in the portion of data selected from the lessons. However, the pre- and post-tests arguably represent the most valid comparison because they offered an identical experience for all participants, whereas the lessons were substantially different between the guided and unguided groups, and not made available to the control group. Furthermore, the pre- and post-tests required participants to transfer their skills to a new topic—namely, the understanding of pattern—which makes a stronger case for the existence of an effect.

Modesty is defined as language that acknowledges uncertainty, whether by making use of conditional forms of verbs, temporal references that suggest things may change, or metacognitive words that reflect the subjectivity of personal observations and the conclusions that one may draw from them. As with interpretation, there was a significant increase in the use of modest language by participants in the guided condition only, with no significant change in either the unguided or control groups. It follows that the presence of modest language correlates relatively strongly with the use of interpretations in the guided condition, $r(18) = 0.61$, $p < .01$. In other words, interpretations may create opportunities for participants to use modest language.

It is not possible from this study to determine what specific factors or combination of factors led to these results. The guided condition was conceived as a holistic experience, with video segmentation as one important component of several; in addition to segmentation, it seems likely that the questions and sample responses that followed each clip also contributed meaningfully to the participants' experience with these lessons. The clips were carefully selected to reflect decision points and opportunities for interpretation, to point out to participants what was important for them to notice as well as offer them an opportunity to stop and think about what they had just observed, and even review the video again, as many times as they wanted. Even if video segmentation had been the only feature of the guided video condition, perhaps it would have been sufficient to draw participants' attention to the content and lead them to closer viewing.

Each clip was followed by a question designed to require an interpretation, ranging in complexity from simply describing what was observed ("he counted using his fingers") to offering one or more theories about what the child understood ("his use of fingers might confuse his counting, which was actually much more competent when the interviewer supported his counting by using her own fingers"). The questions were worded to encourage participants to think carefully and to avoid suggesting a correct answer. However, it is also possible that more generic or less thoughtful questions would have led to a similar degree of close viewing.

The "sample student responses" that followed participants' submitted responses served a dual purpose: first, to help participants gauge their own response by comparing it to one or more responses that were offered as "plausible" and not necessarily authoritative or comprehensive, and, second, to model the language of intellectual modesty that the participants in the guided condition would ideally imitate in their own writing. It seems likely that the sample responses played an important role in the guided group's development of modesty, probably through participants' repeated exposure to the language leading to their gradual adoption of the style in which the feedback was written.

In spite of the guided group's improvements in interpretations and modesty between the pre- and post-tests, there were no significant changes detected for either variable among the five lessons, whether examined in the order in which they were completed or by lesson type. The lack of any finding countered the expectation that participants in the guided group would demonstrate a steady increase in performance on both variables, or perhaps a sudden increase when they adopted an orientation toward interpretation or modesty. One hypothesis for this apparent lack of change is that participants in the guided video group had already responded to as many as 20 questions per lesson by the time they reached the summative questions, and they may have been too fatigued provide as robust a response to these questions as they did in the pre- and post-tests. Alternatively, perhaps they were simply less inclined to repeat what they had already written in response to the preceding questions, which cover the video in great detail. The full text of participants' responses to the guided lesson questions was not analyzed for the purposes of this study, so a logical next step would therefore be use the same coding scheme to ascertain whether change happens in a more distributed way across responses.

For the subtraction and equivalence lessons completed by the guided group, there were significant decreases in the use of interpretation and modesty. These lessons feature more conceptually challenging material than the other lessons. Both guided and unguided groups showed a decrease in claims for these two groups, which suggests that participants may have found less material to identify in the clip. However, the unguided group's lack of significant change on the interpretation and modesty variables suggests that it may also be an issue related to the lesson structure rather than the content.

Limitations and future directions

This study represents the first step in a potential series of studies, including analyses on the existing data as well as new variations on the study. Additional coding on the guided group's responses to each question in the lessons may also reveal some form of progress on the two variables that were significant in the comparison of pre- and post-test responses. Further research that might be conducted on the current data set includes an assessment of how well participants learned the mathematics education content conveyed in the lessons, which was the explicit intention of the lessons even though the present study focused on participants' development of interpretation skills and intellectual modesty. A transfer test could be created for each mathematical topic, using a new video clip similar to the one analyzed by participants for the lesson, and requiring participants to assess the child's understanding, explain the mathematical principles, and develop a strategy for teaching. This type of learning assessment should seek to determine whether the participants who improve their interpretation and modesty also learn the content more effectively.

If the study were to be revised and repeated, it would be useful to add quantitative measures to the lessons and to the pre- and post-tests to measure participants' understanding of the content, their assessment of the child's understanding, their confidence in these assessments, and other factors. Quantitative data would provide a balance to the qualitative data and make comparisons within and between subjects easier. It would also be interesting to compare variations on the guided lesson design to test whether segmentation alone, without questions and/or sample responses, can have a similar effect on interpretations or modesty.

The guided video lesson design could also be applied to other domains, whether for teacher education or other fields. Within teacher education, it would be interesting to follow new teachers who had shown improved interpretation skills and intellectual modesty—ideally in a more robust way than demonstrated in the present study—to investigate whether these changes made an impact on teaching, assessment strategies, attitudes toward teaching and assessment, sense of self-efficacy, and understanding of children's learning. Teachers who participated in an unguided lesson version using the same video could serve as a comparison group.

Conclusion

This study tests a video-based method that was developed as part of a larger initiative to create an early mathematics education curriculum and pedagogical approach to meet the challenges outlined by the NAEYC-NCTM joint statement, which emphasizes the need for teachers to possess a psychological understanding of children's mathematical learning and thinking, employ sensitive assessment techniques for determining what children already know and what they need to learn, and use developmentally appropriate teaching methods based on that understanding (NAEYC & NCTM, 2002).

A video-based method of instruction has a number of implications for teacher development. First, it grounds teachers' learning in the empirical. Second, it offers preservice teachers a means of practicing their skills of observation and evidence-based interpretation (based in part on skills of argument) in preparation for entering the classroom. The goal of this method is to instill in teachers sensitivity to evidence and to encourage them to be deliberate in advancing their understanding of children and teaching. Third, the method may also help teachers develop intellectual modesty—that is, to acknowledge the boundary between what they know and what they must investigate further, ultimately to help them make better-informed decisions about their teaching.

Pedagogy matters. The mere availability of video is not sufficient for making a change in certain types of thinking, but segmenting a video and asking reflective questions may help teachers develop a sort of "skeptical practice" in the face of their preconceptions and understandings. The orientation toward interpretation and modesty places teachers in a better position to change their minds when new evidence comes along. Complexity is inherent in assessing and teaching children, and one goal of educating teachers should be to help them embrace this complexity and develop a propensity to examine it more closely and to interpret it with appropriate intellectual modesty.

References

Dewey, J. (1910). *How we think*. New York: Heath.
Edman, L. R. O. (2008). Are they ready yet? Developmental issues in teaching thinking. In D. S. Dunn, J. S. Halonen & R. A. Smith (Eds.), *Teaching Critical Thinking in Psychology: A Handbook of Best Practices*. West Sussex, UK: Wiley-Blackwell.
Eisner, E. W. (1998). *The englightened eye: Qualitative inquiry and the enhancement of educational practice*. Upper Saddle River, NJ: Merrill.
Ennis, R. H. (1987). A taxonomy of critical thinking dispositions and abilities. In J. B. Baron & R. J. Sternberg (Eds.), *Teaching thinking skills: Theory and practice* (pp. 9-26). New York: W. H. Freeman.
Ericsson, K. A., & Charness, N. (1994). Expert performance. *American Psychologist, 49*, 725–747.
Felton, M., & Kuhn, D. (2001). The development of argumentive discourse skills. *Discourse Processes, 32*, 135-153.
Fuller, F. F., & Manning, B. A. (1973). Self-Confrontation Reviewed: A Conceptualization for Video Playback in Teacher Education. *Review of Educational Research, 43*(4), 469-528.
Garson, J. (2009). Modal Logic. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy (Winter 2009 Edition)*. Stanford: Stanford University.
Ginsburg, H. P. (1989). *Children's arithmetic* (Second ed.). Austin, TX: Pro-Ed.
Ginsburg, H. P., Cami, A., & Preston, M. D. (2009). Inquiry practices: How can they be taught well? In N. Lyons (Ed.), *Handbook of Reflection and Reflective Inquiry: Mapping a Way of Knowing for Professional Reflective Inquiry* (pp. 453-472). New York: Springer.
Glaser, E. M. (1985). Critical Thinking: Educating for Responsible Citizenship in a Democracy. *National Forum: Phi Kappa Phi Journal, 65*(1), 24-27.
Halpern, D. F. (1998). Teaching critical thinking for transfer across domains. *American Psychologist, 53*(4), 449-455.
Jacobs, V. R., Lamb, L. L. C., & Philipp, R. A. (2009). Professional Noticing of Children's Mathematical Thinking. *Journal for Research in Mathematics Education, 41*(2), 169-202.
Kuhn, D. (1999). A developmental model of critical thinking. *Educational Researcher, 28*(2), 16-25.
Kuhn, D. (2001). How Do People Know? *Psychological Science, 12*(1), 1-8.
Kuhn, D., Shaw, V., & Felton, M. (1997). Effects of Dyadic Interaction on Argumentive Reasoning. *Cognition and Instruction, 15*(3), 287-315.
Loos, E. E., Anderson, S., Day, D. H., Jordan, P. C., & Wingate, J. D. (Eds.). (2003) Glossary of lingustic terms. Dallas: SIL International.
Miller, K. F., & Zhou, X. (2007). Learning from classroom video: What makes it compelling and what makes it hard. In R. Goldman, R. Pea, B. Barron & S. Derry (Eds.), *Video research in the learning sciences* (pp. 321-334). Mahwah, NJ: Erlbaum.
NCTM. (2000). Principles and Standards for School Mathematics. Reston, VA: National Council of Teachers of Mathematics.

NRC. (2001). *Adding It Up: Helping Children Learn Mathematics*. Washington, DC: National Academies Press.
Paul, R. W. (1995). *Critical thinking: How to prepare students for a rapidly changing world*. Santa Rosa, CA: Foundation for Critical Thinking.
Piaget, J. (1976). *The child's conception of the world*. Totowa, NJ: Littlefield, Adams.
Santagata, R., Zannoni, C., & Stigler, J. W. (2007). The role of lesson analysis in pre-service teacher education: an empirical investigation of teacher learning from a virtual video-based field experience. *Journal of Mathematics Teacher Education, 10*(2), 123-140.
Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction, 16*(4), 475-422.
Schwartz, D. L., & Hartman, K. (2007). It is not television anymore: Designing digital video for learning and assessment. In R. P. R. Goldman, B. Barron & S. J. Derry (Ed.), *Video research in the learning sciences*. Mahwah, NJ: Laurence Erlbaum Associates.
Sherin, M. G., & Han, S. (2004). Teacher learning in the context of a video club. *Teaching and Teacher Education, 20*, 163-183.
Sherin, M. G., & van Es, E. A. (2005). Using video to support teachers' ability to notice classroom interactions. *Journal of Technology and Teacher Education, 13*(3), 475-491.
Star, J. R., & Strickland, S. K. (2008). Learning to observe: using video to improve preservice mathematics teachers' ability to notice *journal of Mathematics Teacher Education, 11*(2), 107-125.
van Es, E. A. (2008). *A framework for learning to notice student thinking*. Paper presented at the American Educational Research Association.
van Es, E. A., & Sherin, M. G. (2002). Learning to notice: Scaffolding new teachers' interpretations of classroom interactions. *Journal of Technology and Teacher Education, 10*(4), 571-596.
van Gelder, T. (2005). Teaching Critical Thinking: Some Lessons from Cognitive Science. *College Teaching, 53*(1), 41-46.