2009-07-01

Supporting Induction Teachers' Development Using Performance-Based Video Evidence

Peter Rich
pjrich@gmail.com

Richard E. West
rw@byu.edu

Michael Hannafin

Arthur Recesso

Craig Shepherd

Supporting Induction Teachers' Development Using Performance-Based Video Evidence

Richard E. West (Brigham Young University)

Peter Rich (Brigham Young University)

Craig Shepherd (University of Wyoming)

Arthur Recesso (University of Georgia)

Michael Hannafin (University of Georgia)

Supporting Induction Teachers' Development Using Performance-Based Video Evidence

Abstract

This formative research study was designed to inform the development of the Teacher

Success Model (TSM), an initiative to develop a systematic, evidence-informed model for

teacher assessment. While the overall initiative includes all teachers, support professionals,

and evaluators, this study focused on induction teachers and their mentors/administrators.

Participants were shown sample digital videos corresponding to seven pre-selected TSM

attributes, and were asked to identify, annotate, and rate instances where specific attributes

were evident. Overall, our findings indicated that participants could identify examples of key

teaching attributes in the videos and that group reflection and discussion enabled many to

refine their understanding of evidence in assessing teaching. These findings will guide

further development of the TSM assessment model as a multi-faceted, evidence-informed

teaching assessment model.

Supporting and Assessing Induction Teachers' Development

Using Performance-Based Video Evidence

Researchers in the United States have long advocated support for beginning teachers (e.g., Smith & Ingersoll, 2004). As a result, support programs for induction teachers, those with less than five years of experience, are available in roughly 80% of schools and districts (Fulton, Yoon, & Lee, 2005). Typically, induction programs include orientations, retreats, peer groups, mentors, workshops, action research, teamwork, and self-study with the goal of promoting professional development and retention. Often, these programs help teachers to examine their teaching, explore alternative pedagogical strategies, and assimilate local policies (Kelchtermans & Ballet, 2002; Kooy, 2006; Thompson, 2002). During the induction period, building-level administrators also evaluate teacher practices for recertification, tenure, and promotion purposes, yet this assessment may sometimes undermine efforts to assist teachers when feedback focuses on general practices as opposed to specific improvements (National Research Council, 1996; Shymansky & Kyle, 1992).

To strengthen induction support and assessment, our research team is developing tools and procedures for assisting teachers, mentors, and administrators in linking evidence of individual teaching practices to attributes in a teaching framework. In this article, we review the findings from previous research related to induction practices, and present the Teacher Success Model (TSM) framework. We then report findings from a formative research study during which induction teachers, mentors, and building-level supervisors identified and interpreted teaching attributes using the Video Analysis Tool (VAT)—a technology used to identify, annotate, and rate video evidence using rubrics of teaching attributes. We conclude by providing cautions and

recommendations for implementing similar approaches for induction teacher assessment and support.

Supporting Induction Teachers

Induction teachers report being often overwhelmed by the challenges of learning curriculum, managing classrooms, refining teaching methods, and integrating into school, faculty, and community cultures (McCaughtry, Kulinna, Cothran, Martin, & Faust, 2005). Mentors provide support and encouragement to novices as they learn to cope with these responsibilities (Carver & Katz, 2004; Feiman-Nemser, 2001b). Kelley (2004) and Smith (2007) reported that beginning teachers who were mentored were less likely to leave the field or transfer from their schools, and Harrison, Lawson, and Wortley (2005) found novice teachers who were mentored developed professional independence and became critically reflective.

Yet, implementing induction and mentoring programs has been problematic. Such programs often lack the systematic rigor needed to differentiate assistance based on their unique needs (Garet, Porter, Desimone, Birman, & Yoon, 2001). Induction programs also vary among schools, districts, and states within the U.S. Some span multiple years and include orientations, workshops, one-on-one mentoring experiences, and related professional development activities (Carver & Katz, 2004), but other programs offer little formal support (Kozol, 2007). According to Garet et al's. (2001) survey of 1,027 mathematics and science teachers, the most effective professional development programs were sustained over time (as opposed to isolated workshops) and provided opportunities to both observe expert teachers and to be observed.

However, providing such support has proven difficult. Novice teachers typically lack pedagogical knowledge, mentors may espouse competing or contradictory teaching philosophies and beliefs, and teaching schedules often interfere with planned observations. Johnson (2004)

reported that mentors who teach in different grade-levels, subject areas, and schools struggle to provide authoritative and relevant advice (Carver & Katz, 2004). McCann and Johanessen (2004) and Bullough et al. (2008) argued that little time and energy remains after addressing daily classroom needs, and induction teachers consequently often elect not to participate in teaching communities or mentoring programs. Whereas mentoring is considered most effective when mentors and mentees collaborate, try novel teaching approaches, and jointly explore teaching practices (Harrison, Dymoke, & Pell, 2006; Howe, 2006), but even when mentoring is provided, mentors often focus instead on emotional rather than on professional support (Carver & Katz, 2004; Feiman-Nemser, 2001b).

Assessing Induction Teachers

Besides supporting induction teachers, administrators are responsible for ensuring that students' learning is not compromised. The United States' No Child Left Behind (NCLB) legislation (107th Congress, 2002) required that teacher and school assessments be based on student performance outcomes (standardized test scores)—a policy many have argued lacks validity (see, for example, Garfield, 1994; Wolf, Bixby, Glenn, & Gardner, 1991). For example, few well-defined and validated benchmarks have emerged for assessing induction teacher practices and for gauging teacher growth over time. The increased emphasis on national standards and alignment between expected outcomes with student performance have been addressed to some extent [e.g., National Council of Teachers of Mathematics (NCTM), National Center for Science Education (NCSE), National Board for Professional Teaching Standards (NBTS)], but benchmarks and tools are needed for documenting, monitoring, and facilitating teacher progress in associated teaching practices.

Feiman-Nemser (2001a) advocated both supporting *and* assessing new teachers' growth based on tangible evidence of effectiveness. Student outcomes provide one indication of teacher quality (Darling-Hammond, 2002; Goldhaber & Brewer, 1997), but often do little to clarify the practices associated with those outcomes. Our research collaborative is creating a model to support teacher development using multiple sources of evidence (Authors, 2008b, in press-a). Paradoxically, the need appears greatest where support is weakest. Disadvantaged schools lack the resources, tools, and access needed to provide the kinds of assessments likely to promote induction teacher development (Hurd, 1997; Jinks & Lord, 1990; Luft & Cox, 2001; Luft, Roehrig, & Patterson, 2003). Technology-based assessment and support models might provide a solution, as it can enable teachers to connect to mentors in their subject area through the Internet outside of class.

The Emergence of Video in Assessing and Supporting Teacher Development

Recently, researchers have employed digital video for facilitating mentoring and professional development through asynchronous observation (e.g., Authors, 2008d, in press-b, Miller & Glover, 2007; Sherin & van Es, 2005). Van Es and Sherin (2008) found that when mathematics teachers presented and analyzed short video segments of each other's practices, they examined their own practices differently, became more nuanced in the classroom phenomena they discussed, used multiple approaches to analysis, and reconceptualized their conceptions of teaching. Richert (1990) suggested that teaching artifacts facilitated reflection on practice because they re-created phenomena and oriented teachers to specific classroom attributes. According to van Es and Sherin (2008), video-based practices may increase the accuracy of classroom re-creations and provide an impetus for expert and novice review.

Thus far, research has not conclusively documented the accuracy with which induction teachers can assess their practices using video. During initial student teaching field experiences, for example, Authors (2006) found that while beginning teachers became more nuanced in their examinations of classroom practices through sustained video analyses, they missed important insights detected by expert teachers. Thus, induction teachers may require mentor or supervisor support while analyzing their video-recorded practices. The purpose of this study was to examine how induction teachers, mentors, and administrators used a video tool to identify and rate specific teaching attributes, and how the discussion about the video evidence influenced the participants' (especially the induction teachers') perceptions of teaching practices.

Research Questions

We posed two research questions:

1. How do induction teachers, mentors, and leaders identify specific attributes of teacher practices in video recordings of classroom teaching?

2. How does the use of video evidence influence how teachers, mentors, and leaders discuss and reflect on teachers' practices?

Methods

Participants

Induction teachers, as well as their mentors and administrators, from an urban high school, rural middle school, and rural elementary school were purposefully selected (Merriam, 1998; Patton, 2001) to represent different school grade levels and different areas of the state. Twenty-six teachers, mentors, and administrators agreed to participate, including 12 induction teachers (less than 5 years teaching experience), 7 mentor teachers, and 7 principals or vice-principals.

The Teacher Success Model (TSM)

The Teacher Success Model is a framework with tools and processes for identifying and analyzing teachers' performances in critical teaching domains. While this study piloted a small aspect of the model, that of video evidence for assessing only seven teaching attributes, the initiative's overall goal is to generate detailed descriptions of many different teaching attributes along with indicators of corresponding, representative kinds of evidence that embody the attributes in practice. Using these descriptions, we theorize that participants can focus on specific teaching practices without being distracted by extraneous or unrelated elements of the teaching experience, in much the same way that photography lenses amplify or suppress various attributes of an image. Our collaborative seeks to understand whether the panorama of teaching can be better understood by temporarily and systematically focusing on these specific attributes.

To develop the framework, we first synthesized the literature on developmental stages of teacher expertise (e.g., Alexander, 2003; Berliner; 1986, 1994, 2004; Bond, Smith, Baker, & Hattie, 2000; Dunphy & Williamson, 2004; Feiman-Nemser, 2001a; Fuller, 1969). Next, we adopted six domains the existing state framework for accomplished teaching (Georgia Systemic Teacher Education Program [GSTEP], 2003). GSTEP domains included planning and instruction, assessment, knowledge of students and their learning, learning environment, content and curriculum, and professionalism. Working within the GSTEP framework, we identified attributes of teachers' practices within each domain and created a continuum for teacher development that ranged from basic to advanced practice. Based on published teacher development literature, we then identified performance indicators within each attribute and developmental level. Throughout this process, experts from the state department of education, former teachers and administrators, and educational researchers vetted these attributes,

continuum levels, and indicators. Figure 1 shows an example of one attribute in the TSM framework. In this study, seven attributes were used as a lens for assessing teachers' performances. The complete TSM framework and rubric is available at: http://lpsl.coe.uga.edu/tsm.pdf [Note the continuum levels (basic, proficient, and advanced) used in this study were subsequently revised (emerging, refining, and excelling)]. We are currently identifying appropriate forms of evidence for each attribute and developing processes for administering the framework. Ultimately, we hope the TSM will enable induction teachers and support personnel to assess progress toward a range of teaching benchmarks and focus and customize support for development in specific areas.

Because teachers develop uniquely, the TSM continuum represents progressive refinements of specific teaching practices and not an overall evaluation of the teacher. In other words, teachers may refine certain aspects of their practice while remaining at a basic level in others. This differentiation in performance supports targeted professional development and mentoring while avoiding an overall label that could stigmatize teacher performance.

Video Analysis Tool

The Video Analysis Tool (VAT) was employed to isolate video clips from larger digital video files, categorize and rate these clips according to the TSM rubric, associate comments with the clips, and share the clips for feedback and assessment purposes (Authors, 2008a, in press-a). Although teachers have long used video evidence for assessment and development purposes (Barber, 1990; Fuller, 1969; National Board for Professional Teaching Standards, 2004), previous technology has been difficult to implement. As shown in Figure 2, the VAT enables a user to view, assess, annotate, and link evidence with teaching events (Authors, 2008b, 2008c).

The TSM lens shown was given to the participants on paper because the rubric prototype had not yet been incorporated into the Video Analysis Tool until after this study.

Procedures

We presented the Teacher Success Model to participants at each school and asked them to identify evidence in sample digital videos corresponding to seven pre-selected TSM attributes (see Table 1). These attributes, selected from a pool of roughly 40 TSM attributes, were considered most likely to be observable in everyday teacher practice. We chose this method of attribute sampling because this was a formative study designed to observe how participants identify and interpret common teaching attributes based on an assessment rubric. Prior to analysis, participants indicated that they understood and believed they could recognize the attributes in typical classroom settings and successfully use the VAT.

Teachers recorded themselves teaching two lessons, and researchers imported the digital videos to the VAT. Participants used the VAT and the rubric of sample TSM teaching attributes to view and annotate videos of the induction teachers' practices. Induction teachers annotated their own videos, while mentor teachers and administrators annotated each induction teacher's videos. When annotating, we asked participants to divide videos into segments showing the specific instances of the attributes in practice, to rate the practices using the corresponding TSM rubrics (basic, proficient, advanced), and to write comments about the clips. During their first analysis, many participants marked each instance of a TSM attribute as requested, while others marked only the first instance noticed. After observing this initial confusion, we provided additional one-hour training to overcome technological hurdles and clarify the need to identify every noticed instance of a practice. Participants' assessments during their subsequent analyses then improved.

Data Collection

We conducted two focus group interviews at each school, focusing on participants'

experiences using the VAT and the TSM guidelines for identifying key teaching attributes. The

first interview occurred following initial annotation, and focused on experiences in using the

VAT, the attributes of teacher practices, and the TSM continuum. We asked participants which

attributes they could readily identify in the videos and why particular attributes may or may not

be observable using videos of teacher performance. We also facilitated discussions between the

mentors and teachers about what constituted evidence of particular teaching attributes. After the

participants annotated their second video recordings of each induction teacher practice, we

conducted a second focus group interview using the same protocol. Additionally, we probed to

identify differences in participants' experiences from the first set of videos to the second and

asked what their perceptions were of the utility of reviewing teacher practices through the VAT.

Data Analysis

Using principles of grounded theory (Glaser & Strauss, 1967; Strauss & Corbin, 1998),

three researchers independently coded the focus group transcripts, overlapping so that each

transcript was coded twice by different researchers. After initial coding, the researchers

compared codes and identified possible themes and categories for the segments, noting emergent

patterns within and across cases. The resultant thematic hierarchy provided the structure for

coding the remaining data (See Table 2). We used constant comparison techniques as we coded

to continually refine the themes.

<div align="center">Findings: Identification of Teaching Practices</div>

Two major themes emerged about how induction teachers, mentors, and administrators

identified attributes of teaching using videos of teacher practice. First, participants differed in

their abilities to discern examples of teaching attributes as a function of their unique teaching backgrounds. Second, participants struggled to classify the examples of teaching attributes along the TSM continuum. They expressed uncertainty in how to differentiate the levels of the continuum, tending to override the TSM structures with their own past assessment practices.

Differences in How Participants Observed Practice

Rural County Elementary School. The teachers began the first focus group session noting that they could identify planning attributes in their teaching because, as one teacher said, "when we're watching our own tapes, we understand our own background." This was supported by another induction teacher that explained, "If it's your class, you can see why you asked someone a certain question." A mentor noted that planning would, "definitely be easier to see with a lesson plan," while another mentor remarked, "You just about always have to ask the person why they did something a particular manner." Thus, all teachers reported that personal experience in the situation was central to judging the quality of instructional planning.

In contrast, teachers noted the need for content-knowledge experience for non-planning attributes, such as whether or not the lesson met grade-level standards. A fifth-grade mentor teacher noted of her induction teacher's lesson, "it was very obvious that the teacher was prepared for the lesson." The induction teacher explained, "that [a person reviewing the video] must be from the same grade level to be able to see standards." This notion was supported by the administrator as well as a second-grade mentor teacher that indicated she felt less prepared to judge the fifth-grade standards.

During the second focus group interview, teachers again noted the importance of personal experience and content expertise when analyzing videos. In addition, the importance of teaching expertise emerged during the second interviews. When discussing techniques for differentiation,

one induction teacher mentioned being unable to recognize techniques in her own video. However, two mentors identified a specific example of that teacher's effective differentiation. The induction teacher remarked, "I wouldn't have even thought of the vocab game as being any of the attributes or levels. . . . I never thought of that as a different strategy." Thus, participants were able to discern attributes of teacher practice differently based on their teaching expertise.

Rural County Middle School. Diverse teaching and professional backgrounds also influenced how these participants identified teaching attributes in their videos. For example, experienced administrators and mentors helped to identify attributes that induction teachers could not. At one point, two teachers debated whether it was impossible to identify rapport with students in a video, since rapport is not directly observable. However, a principal noted an instance when a student understood the teacher's comment as a joke and was not offended. With a different induction teacher, an administrator identified an instance where the teacher laughed *with* her students, demonstrating rapport that she had not considered. During the discussion, one teacher suggested that it was impossible to observe evidence of the teacher's content knowledge in a video: "If you want to analyze my content ability, look at my praxis scores." However, his mentor teacher helped him identify video evidence for content knowledge by explaining, "When you stop and elaborate with students, that shows your content knowledge." After discussing how the more experienced participants observed this attribute, the teacher reported during the second focus group interview, "last week this was probably the hardest one for me to find [but] I had an easier time finding it this time." Thus, it appears that the discussion with the more experienced teachers may have helped the induction teacher to better identify evidence for this attribute.

When examining attributes related to learning environments, participants reported that the teachers' knowledge of students was more important than an administrator's or mentor's

expertise. One teacher, for example, described using different color handouts because they helped a student, and suggested that an outside evaluator might simply conclude the teacher liked green paper. She also explained that not finding visible evidence of class management might indicate advanced teaching practices because the teacher is maintaining control via subtle nods, glances, or carefully chosen words to correct students. "Part of being good with a student," she said, "Is that the teacher doesn't want anyone else to know she is disciplining them." Thus, rural middle school participants reported that, depending on the domain being assessed, it was important to either have intimate, first-hand knowledge of the students and the situation or to have extended experience as a teacher.

Urban City High School. Induction teachers reported the fewest difficulties identifying attributes, seemingly because they developed the lessons and knew their students. The easiest attributes for participants to identify were respect and rapport for students and classroom management. One induction teacher said

> At one point I was explaining something to them on the board and I could hear two students talking . . . and I said to them 'guys.' I just looked at them and there was dead silence . . . Obviously there is a level of rapport, a level of respect with them. I didn't have to yell at them, I didn't have to raise my voice.

This same teacher easily identified other instances in her video. Similar comments were typical across participants, and they felt that video footage easily captured classroom setup, order, activities, and whether or not students were on task.

However, identifying evidence of content knowledge and individual needs proved more difficult, especially for administrators. The principal mentioned, "Things like rapport and classroom management, that's easy to see. But. . . I don't know math. It would be difficult for me

to say." A mathematics mentor made similar comments when viewing science lessons: "She [the science teacher] could have said anything, and I would have said 'well I guess that sounds good to me'." This mentor also wondered, "how much [of] the content area can really be observed" if mentors have different content expertise than their protégés. A mathematics teacher agreed noting "I could say that 2+2 was 6 and do a good job of it, but that wouldn't be correct." Citing these difficulties, one vice-principal mentioned that when evaluating, she focuses on pedagogy instead of content knowledge: "I am a home economics person and an administrator . . . Because of ethics [the teacher] is supposed to be teaching her subject. But . . . I'm watching for how she's doing it." The administrator further stated that observers often assume that teachers present information accurately, and argued for other evaluations to detect content inaccuracies.

Struggles to Rate and Qualify Practice

Rural County Elementary School. The TSM is designed to identify and assess specific teaching practices as they emerge, not provide an overall rating of teacher quality. However, participants expressed confusion in using TSM ratings, tending to override the independent criteria for each attribute and fixating instead on what they assumed to be the most appropriate level, overall, for an induction teacher. For example, one induction teacher reported that she "didn't go and look at the advanced and the other areas." Consequently, we asked participants to apply the entire continuum during their second video analysis. On the whole, participant ratings became more diverse. One induction teacher reported, "after [the] second try I found differences between domains to be much clearer." A mentor declared that she, "flew through the second set because [she] knew the rubric better." Other mentors and the administrator agreed. Once participants understood that practice could be evident anywhere on the continuum, their enthusiasm for teacher assessment improved. As one participant noted:

I think it wouldn't be as scary to add a 'basic' on the continuum because it is just one little element of practice at one point in time. . . . this rubric [in this study] is a lot easier because it's not so negative [as the current state assessment process]. It tells you what to do to get better.

Rural County Middle School. Even after clarification by the researchers, the participants struggled to conceptualize the continuum as representing growth rather than focused targets for improvement. One administrator commented that a teacher attribute "takes years to master" and thus could not be observed with induction teachers. This administrator did not recognize that "basic" indicators provided novice examples of the attribute. Participants reported that it was more difficult to qualify attributes on the continuum than to simply identify them, and it was difficult to do both concurrently. One administrator commented, "If we had to do all that too, the process gets a whole lot harder." However, an administrator remarked, "The more times you do it, the easier it becomes."

Urban County High School. Although participants demonstrated little difficulty locating classroom practices along the rubric continuum, they often struggled when differentiating between levels on the continuum and understanding and applying definitions. A mathematics teacher stated "the wording is so similar . . . I had a hard time defending if I called [a specific practice] advanced or if I called it proficient." A science teacher agreed: "There should definitely [be] a specific checklist . . . This should be basic, this should be advanced, this should be exemplary. Then we could say, 'yes, I saw this' or 'No, I didn't see this.'" Participants agreed that qualifying the attributes along the continuum was difficult due to ambiguities in wording and level descriptions.

Findings: The Impact and Utility of Digital Video Evidence

Since participants were not accustomed to using digital video evidence of teaching performances, they noted both positive and negative effects for the use of video evidence of teaching practices. These effects were identified for two areas: The nature and limitations of the video medium, and the practicality of using video in teacher assessment.

The Nature and Limitations of Video Evidence

Rural County Elementary School. Participants unanimously agreed that video could not adequately capture all instances of practice. All participants reported that differentiated instruction was difficult because, as one induction teacher said, "the reason someone did something different is in the teacher's head." Video did not reflect out-of-classroom activities, as an induction teacher noted: "Seeking support from others is not possible to catch on camera." A mentor agreed, adding that "one thing that you can't see [on camera] are the meetings in which teachers make the choices about how to address an individual student's needs." Still another induction teacher identified the need for "multiple classes to have evidence. . . . [Change] happens over a little bit of time." Thus, while video helped participants to identify some practices, other practices required different types of evidence.

Rural County Middle School. Participants indicated that video evidence was useful, but limited. Participants suggested that video may provide a limited and potentially misleading representation because it is often focused on the teacher's actions during a lesson. As one teacher said, "Most of [the] time, in that particular class, it's not even about me at all . . . I'm not doing anything. They're doing their own learning." Others stated that video is not sufficiently precise to detect subtle teacher actions. One teacher said, " I can be tasking five different things, and if you don't see that whole classroom, you won't know I'm in control of those five different things."

Participants also emphasized that video evidence comprises only a piece of a complex puzzle ("You have to take it as a whole picture, and you're not going to get [it in] a clip") and that attaching lesson plans and other materials to a video clip would be more useful: "How [else] could someone . . . know what [that] the teacher is doing is aimed at suiting different students' needs?"

Urban County High School. Similarly, high school participants mentioned that some TSM indicators were not observable in video recordings and that lesson plans, student work samples, and other materials were necessary to make accurate judgments. For example, several basic teaching attributes implied required support. One induction science teacher argued, "A lot of the support that you would need cannot be captured during the time that you're actually in the classroom." However, participants also noted benefits for using video to encourage discussion. Through video-enhanced discussions, participants explored classroom practices and voiced their opinions. When problems appeared, groups addressed them and brainstormed solutions. These conversations simplified the second round of video analysis. One mentor teacher stated, "I think it gets easier the more you watch . . . As I watched the videos, I learned the attributes. As I learn[ed] the attributes, I notice[ed] them quicker and more often." With continued practice, attribute identification became clearer.

The Practicality of Video Assessment

Rural County Elementary School. Participants largely agreed that the video analysis was useful for teacher assessment. A mentor commented, "I think the ability to watch myself on video is a valuable way to look at your teaching. It lets you set goals for what you need to work on." An administrator also agreed, but despite the utility of video for formative development, one induction teacher reiterated that video was more powerful in combination with other evidence: "I

can't say watching live or watching video is better. You've got to be able to see both." To improve the video analysis process, participants cited the importance of collaboration: "I would have got a lot more out of it if I could have sat next to another teacher."

Rural County Middle School. Participants also reported that the video analysis was useful for reflection and assessment when combined with other data. One teacher said, "It's always good to go back and look at yourself, because it's easy to forget what you do in the classroom." He added that it would be more effective to share video comments with each other: "I would like to be able to see what others said and see their video, too." In addition, the teachers valued the control afforded via video analysis because they could determine which lessons to record. Finally, administrators noted that video-enhanced, TSM-based assessment was more productive than current observation practices that rely on observing classes and completing checklists of teacher qualities because observers can rewind and review the video and record more complete and accurate comments than is possible during real-time observations. An administrator added that observing a teacher through video could help to scaffold the development of inexperienced administrators: "Having this system will probably do the teachers a better service . . . because of my lack of experience."

Urban County High School. Although participants stated that video analysis was beneficial, they recommended additional support structures in future implementations. During the first focus group, the mentor stated, "The further I got along, the more I started questioning myself. 'Am I really seeing this? Is this what this really is?" After initial training, the participants requested additional feedback and support during video analyses in the form of example portfolios, example practices, and definitions of the continuum. Later, participants gained

confidence in their abilities to identify evidence of teaching practices in the videos after discussing and working through their questions as a group.

General Discussion

Overall, our findings suggested that teachers, mentors, and administrators successfully detected examples of key teaching attributes in the videos, although some attributes were more easily identifiable to some groups of teachers because of their background and experience. Similarly, Sherin and van Es (2005) conducted several studies into the use of video clubs for teacher assessment and reflection. In one study, they asked middle-school mathematics teachers and high school mathematics and science teachers to meet monthly to discuss excerpts from their videos. These teachers improved over time in their abilities to extract and interpret relevant components from videos.

Consistent with previous research, the current study also suggested that group reflection and discussion using the videos enabled participants to clarify definitions of teaching practice, verbalize and identify their own biases, and learn from others' perspectives on successful teaching. For example, Thomson (1992) reported that teacher-supervisor collaborative discussions using videos of teacher practice resulted in optimum professional learning for the teacher. Others have documented beneficial effects for teachers who discussed videos of their practice with peers (Sharpe et al., 2003).

It may be especially important for teachers to involve others during video review and discussion in order to benefit from diverse perspectives, expertise, and experience. Sherin (2003) documented how the involvement of a university researcher with six high school mathematics video clubs improved teacher reflection due to the varying expertise available. This varied expertise deepened the discussion of the videos, which helped modify participants'

interpretations of the teaching practices and aided them in engaging in pedagogical change. More recently, van Es and Sherin (2008) described a model for how mathematics teachers in video clubs develop their abilities to notice and interpret students' actions caught on video. They found that teachers developed their professional insights along direct, cyclical, and incremental paths as they participated.

Frederiksen, Sipusic, Sherin, and Wolfe (1998) wrote that professional practice improves when teachers develop a socially shared language of practice for describing attributes of teaching, and then use this language in discussing their videos with peers. MacLean and White (2007) found that preservice and experienced teachers jointly constructed professional teacher identities through discussions of their observations of teaching videos. These studies indicate a need for teachers to collaboratively discuss, assess, and reflect on their practices, and video evidence can structure and facilitate these discussions.

Our findings also reinforced the importance of basing assessments of teachers' practices on multiple rather than singular pieces of evidence. TSM is designed to document developmental combinations of specific actions and intentions that can be observed for supporting both teacher assessment and development. In this study, participants were only asked to do the first part of the process—observe and interpret specific actions. This proved difficult, as participants identified numerous other kinds of evidence needed to interpret the videos, such as lesson plans, teacher explanations and reflections, and contextual knowledge of the students and school. One teacher, voicing the opinions of most participants, said, "If I attach my classroom philosophy or plan to the videos, then you might see how that influences the students and their needs and my needs and the school and how that all fits together." The need to associate multiple pieces of evidence is similar to research by Anderson, Schum, and Twining (2005) who stated that no single piece of

evidence can re-create practice because intent, outcome, and other elements may not be directly observable. In order to overcome these limitations, they suggested capturing multiple sources of evidence to create more nuanced depictions of events. In future iterations, we will be attempting to balance the benefits of examining specific details of an experience with converging sources of related evidence.

Participants in this study highlighted important limitations in both the methods and uses of video. While researchers have lauded video as being a reliable and credible type of evidence because it unobtrusively captures what happens in a classroom (Frederiksen, et al., 1998), it may also inherently limit the assessment perspective. Through discussions with participants, it became clear that the robustness of video evidence is influenced by camera location, angle and viewpoint, and the specific events captured. Video evidence may fail to detect nuances such as non-verbal communication and fine-grained visual and auditory cues central to interpreting teacher action. Berg and Smith (1996) noted similar limitations to video evidence and reflection, and Sherin and van Es (2006) found that video evidence forced viewers to examine practices from limited vantage points (a "keyhole effect," much like viewing a classroom through a keyhole). While this limited observation potential, it also forced teachers in Sherin's study to plan collection strategies specifically to inform practices.

<div align="center">Implications and Conclusions</div>

In this study, participants assessed teaching practices, captured in video recordings, using a rubric defining differential expertise. We found that the participants were able to identify examples of teaching attributes in the videos, but to varying degrees. With some attributes, teachers benefited from intimate knowledge of the setting while with other attributes, participants with advanced teaching expertise were most successful. The participants reported it

was easier to identify instances of a teacher attribute than to assess and rate the quality of the teaching practice. The participants also reported positive benefits including an improved understanding of the teaching attributes, evidence of successful teaching, and their own strengths and weaknesses as teachers. These benefits may assist induction teachers as they transition to the teaching profession and refine their skills. The VAT technology could also enable induction teachers to work with mentors outside of class hours, which would be more practical. However, participants noted that video alone was incomplete and cannot capture many aspects of successful teaching without convergent, related evidence.

Several implications are apparent. First, we identified advantages to using the Teacher Success Model as part of a comprehensive induction teacher assessment framework. Participants indicated that it was helpful to use the framework for conducting in-class observations and supporting post-observation analysis and reflection. Teachers can also use the framework as a tool for assessing their own practices and developing evidence-informed professional development plans. By applying the same criteria and methods used by administrators, teachers' self-improvement plans can be more closely aligned with administrative assessments. As Feiman-Nemser (2001a) stated, "New teachers and those responsible for their learning need a defensible basis for deciding what to work toward and some means of determining how they are doing" (p. 1032-1033). Over time, the TSM initiative hopes to equip teachers with refined self-assessment tools and models for formative, personal development that mirror those used in summative teacher assessments by administrators. This study provides a modest, but essential, contribution to this goal.

Finally, these methods may help to guide portfolio assessment practices. Induction teachers often do not understand what constitutes relevant evidence of their teaching practices,

and thus may emphasize less critical elements over arguably more salient evidence. Our findings suggested that scaffolding may help induction teachers to engage in discussions with mentors to identify relevant attributes of their own teaching. Multiple pieces of evidence, including annotated digital video clips, have the potential to increase portfolio usefulness and accuracy.

In future implementations, participants will need to receive additional training in how to assess and rate the quality of teaching practices. Our participants struggled in recognizing that an induction teacher could be very advanced in some teaching attributes, but still basic in others. Next, additional pilot testing is needed to identify which kinds of evidence (video, interviews, test scores, lesson plans, etc.) are most valid for assessing teaching attributes. In future studies, we also hope to study other, more complex, teaching attributes and the process of integrating attributes together to form more complete pictures of teaching expertise. In addition, tools and methods are needed to connect multiple complimentary sources of evidence to best explain particular situations. Consistent with Schum's (1994) guidelines for evidential reasoning, such methods should help to clarify which kinds of evidence are complementary and which pieces of evidence have the greatest force. This should improve decisions about what evidence to collect and how to use it to assess and improve induction teachers' practices.

References

107th U. S. Congress (2002). *Public Law 107–110: The no child left behind act of 2001.* Washington, D. C.: Government Printing Office.

Alexander, P. A. (2003). The development of expertise: The journey from acclimation to proficiency. *Educational Psychologist, 32*(8), 10-14.

Anderson, T., Schum, D., & Twining, W. (2005). *Analysis of evidence* (2nd ed.). Cambridge, MA: Cambridge University Press.

Authors. (2006). Promoting reflection with a Web-based video analysis tool. *Journal of Computing in Teacher Education*, *23*(1), 31-39.

Authors. (2008a). Creating a patchwork quilt for teaching and learning: The use of learning objects in teacher education. In P. Northrup. (Ed.). *Learning objects for instruction: Design and evaluation* (pp. 261-280). Hershey, PA: Idea Group.

Authors. (2008b). Evidential reasoning and decision support in assessment of teacher practice. In T. J. Kowalski & T. J. Lasley (Eds). *Handbook on data-based decision making in education*. (pp. 363-381). New York: Routledge.

Authors. (2008c). An evidential reasoning approach to analysis of teaching practices using a Web-based video analysis tool. In A. L. Tan and Y. J. Lee (Eds). *Science education at the nexus of theory and practice* (pp. 157-178)*.* Taipei, Taiwan: Sense Publishers.

Authors. (2008d). Examining preservice teacher inquiry through video-based, formative assessment e-portfolios. *Journal of Computing in Teacher Education, 25*(1), 31-37.

Authors. (in press-a). Direct evidence and the continuous evolution of teacher practice. In P. Adamy, & N. Milman (Eds.), *Evaluating electronic portfolios in teacher education*. Charlotte, NC: Information Age Publishing.

Authors. (in press-b). Video annotation tools: Technologies to scaffold, structure and transform

    teacher reflection. *Journal of Teacher Education*.

Berg, M. H., & Smith, J. P. (1996). Using videotapes to improve teaching. *Music Educators*

    Journal, *82*(4), 31-37.

Berliner, D. C. (1986). In pursuit of the expert pedagogue. *Educational Researcher, 15*(7),

    5-13.

Berliner, D. C. (1994). Teacher expertise. In T. Husén & T. N. Postlethwaite (Eds.), *The*

    international encyclopedia of education (2nd ed., Vol. 10, pp. 6020-6026). Elmsford,

    NY: Pergamon.

Berliner, D. (2004). Describing the behavior and documenting the accomplishments of expert

    teachers. *Bulletin of Science Technology Society, 24*(3), 200-212.

Barber, L. (1990). Self-assessment. In. J. Milman & L. Darling-Hammond (Eds.), *The new*

    handbook of teacher evaluation: Assessing elementary and secondary school teachers

    (pp. 216-228). Newbury Park, CA: Sage.

Bond, L., Smith, T. W., Baker, W. K., & Hattie, J. A. (2000, September). *The certification*

    system of the National Board for Professional Teaching Standards: A construct validity

    study. Greensboro, NC: Center for Educational Research and Evaluation.

Bullough, R. V. Jr., Young, J. R., Hall, K. M., Draper, R. J., & Smith, L. K. (2008). Cognitive

    complexity, the first year of teaching, and mentoring. *Teaching and Teacher Education,*

    24(7), 1846-1858.

Carver, C. L., & Katz, D. S. (2004). Teaching at the boundary of acceptable practice: What is a

    new teacher mentor to do? *Journal of Teacher Education*, *55*(5), 449-462.

Darling-Hammond, L. (2002). What's at stake in high stakes testing? *Brown University Child & Adolescent Behavior Letter, 18*(1), 1.

Dunphy, B. & Williamson, S. L. (2004). In pursuit of expertise: Toward an educational model for expertise development. *Advances in Health Science Education, 9*(2), 107-127.

Feiman-Nemser, S. (2001a). From preparation to practice: Designing a continuum to strengthen and sustain teaching. *Teachers College Record, 103*(6), 1013-1055.

Feiman-Nemser, S. (2001b). Helping novices learn to teach lessons from an exemplary support teacher. *Journal of Teacher Education*, *52*(1), 17-30.

Frederiksen, J. R., Sipusic, M., Sherin, M., & Wolfe, E. W. (1998). Video portfolio assessment: Creating a framework for viewing the functions of teaching. *Educational Assessment, 5*(4), 225-297.

Fuller, F. F. (1969). Concerns of teachers: A developmental conceptualization. *American Educational Research Journal, 6*(2), 207-226.

Fulton, K., Yoon, I., Lee, C. (2005). *Induction into learning communities*. New York: National Commission on Teaching and America's Future. Washington D.C.: National Commision on Teaching and America's Future.

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, *38*(4), 915-945.

Garfield, J. B. (1994). Beyond testing and grading: Using assessment to improve student learning. *Journal of Statistics Education, 2*(1), 1-11.

Georgia Systemic Teacher Education Program. (2003). GSTEP Framework. Retrieved May 2008, from http://www.coe.uga.edu/gstep/documents/gstep_framework0603.pdf.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Mill Valley, CA: Sociology Press.

Goldhaber, D.D. & Brewer, D.J. (1997). Why don't schools and teachers seem to matter? Assessing the impact of unobservables on educational productivity. *The Journal of Human Resources, 32*(3), 505-523.

Harrison, J., Dymoke, S., & Pell, T. (2006). Mentoring beginning teachers in secondary schools: An analysis of practice. *Teaching and Teacher Education, 22*(8), 1055-1067.

Harrison, J. K., Lawson, T., & Wortley, A. (2005). Mentoring the beginning teacher: Developing professional autonomy through critical reflection on practice. *Reflective Practice, 6*(3), 419-441.

Howe, E. (2006). Exemplary teacher induction: An international review. *Educational Philosophy and Theory, 38*(3), 287-297.

Hurd, P. D. (1997). *Inventing Science Education for the New Millenium.* Ways of Knowing Science Series. Columbia University, NY: Teachers College Press.

Jinks, J., & Lord, R. (1990). The better elementary science teaching project: Overcoming the rural teachers professional and academic isolation. *School Science and Mathematics, 90*(2), 125-133.

Johnson, S. M. (2004). Schools that support new teachers. In S. M. Johnson (Ed.) *Finders and keepers: Helping new teachers survive and thrive in our schools* (pp. 91-118). San Francisco, CA: Jossey-Bass.

Kelchtermans, G. & Ballet, K. (2002). The micropolitics of teacher induction: A narrative-bibliographical study of teacher socialization. *Teaching and Teacher Education, 18*(1)*,* 105-120.

Kelley, L. M. (2004). Why induction matters. *Journal of Teacher Education*, *55*(5), 438-448.

Kooy, M. (2006). The telling stories of novice teachers: Constructing teacher knowledge in book

      clubs, *Teaching and Teacher Education*, *22*(6), 661-674.

Kozol, J. (2007). *Letters to a young teacher*. New York: Crown Publishers.

Luft, J. A., & Cox, W. E., (2001). Investing in our future: A survey of support offered to

      beginning secondary science and mathematics teachers. *Science Educator, 10*(1), 1-9.

Luft, J. A., Roehrig, G. H., & Patterson, N. C. (2003). Contrasting landscapes: A comparison of

      the impact on different induction programs on beginning secondary science teachers'

      practices, beliefs, and experiences. *Journal of Research in Science Teaching, 40*(1), 77-

      97.

Maclean, R., & White, S. (2007). Video reflection and the formation of teacher identity in a team

      of pre-service and experienced teachers. *Reflective Practice, 8*(1), 47-60.

McCann, T. M. & Johannesen, L. R. (2004). Why do new teachers cry? *Clearing House, 77*(4).

      p. 138-145

McCaughtry, N., Kulinna, P. H., Cothran, D. J., Martin, J. J., & Faust, R. (2005). Teachers

      mentoring teachers: A view over time. *Journal of Teaching in Physical Education, 24*(4),

      326-343.

Merriam, S. B. (1998). *Qualitative research and case study applications in education*. San

      Francisco: Jossey-Bass.

Miller, D. & Glover, D. (2007). Into the unknown: The professional development induction

      experience of secondary mathematics teachers using interactive whiteboard technology.

      Learning, Media, and Technology, 32(3), 319-331.

National Board for Professional Teaching Standards. (2004). *Standards & national board certification*. Retrieved December 2005, from http://www.nbpts.org/standards/nbcert.cfm

National Research Council. (1996). *National science education standards*. Washington, DC: National Academy Press.

Patton, M. Q. (2001). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Richert, A. E. (1990). Teaching teachers to reflect: A consideration of programme structure. *Journal of Curriculum Studies, 22*(6), 509-527.

Schum, D. A. (1994). *Evidential foundations of probabilistic reasoning.* New York: Wiley.

Sharpe, L., Hu, C., Crawford, L., Gopinathan, S., Myint, S. K., Moo, S. N., & Wong, A. (2003). Enhancing multipoint desktop video conferencing (MDVC) with lesson video clips: Recent developments in pre-service teaching practice in Singapore. *Teaching and Teacher Education*, *19*(5), 529-541.

Sherin, M. G., & van Es, E. A. (2006). What (we think) we know about video clubs. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.

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.

Sherin, M. G. (2003). Using video clubs to support conversations among teachers and researchers. *Action in Teacher Education, 24*(4), 33-45.

Shymansky, J. A., & Kyle, W. C. Jr. (1992). Establishing a research agenda: Critical issues of science curriculum reform. *Journal of Research in Science Teaching, 29*(8), 749-778.

Smith, T. M., (2007). How do state-level induction and standards-based reform policies affect induction experiences and turnover among new teachers? *American Journal of Education, 113*(2)*,* 273-309.

Smith, T. M. & Ingersoll. (2004). What are the effects of induction and mentoring on beginning teacher turnover? *American Educational Research Journal. 41*(3), 681-714.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Grounded theory procedures and techniques* (2nd ed.). Newbury Park, CA: Sage.

Thompson, P. H. (2002). Data help newcomers analyze their teaching. *Journal of Staff Development*, *23*(4), 57-59.

Thomson, W. S. (1992). *Using videotape as a supplement to traditional student teacher supervision*. ERIC Document Reproduction Service. ED357014.

van Es, E. A. & Sherin, M. G. (2008). Mathematics teachers' "learning to notice" in the context of a video club. *Teaching and Teacher Education*, *24*(2), 244-276.

Wolf, D., Bixby, J., Glenn, J. III, Gardner, H. (1991). To use their minds well: Investigating new forms of student assessment. *Review of Research in Education, 17*(1), 31-74.

Table 1.

TSM Domains and Attributes

Domain 1: Content & Curriculum. Teachers demonstrate a strong knowledge of content area(s) appropriate for their certification levels.
      Attribute A: Content knowledge in teaching area.
      Attribute E: Resources and content area knowledge.
      Attribute F: Curriculum aligned per state and national content area standards.

Domain 2: Knowledge of Students & Their Learning. Teachers support the intellectual, social, physical, and personal development of all students.
      Attribute C: Respect for and rapport with students
      Attribute E: Accommodation of individual student needs

Domain 3: Learning Environments. Teachers create learning environments that encourage positive social interaction, active engagement in learning, and self-motivation.
      Attribute C: Classroom management.
      Attribute E: Individual differences in classroom

Note. The teaching attributes studied in this research represented three domains from the Teacher Success Model. These attributes were selected because they were likely to be identifiable in most teaching episodes.

Table 2.

Coding Categories, Definitions, and Frequencies.

Category (Freq.)	Codes (Freq.)	Definitions
Technical Issues (10)	Camera setup/playback (8)	Issues dealing with the physical setup, placement, and playback of video cameras
	Ease of use (2)	Issues related to technical functionality.
VAT (23)	Ease of use (11)	Personal ease in navigating and manipulating functions in VAT
	Functionality (7)	System providing or not providing the tools to manipulate data in the desired or expected way
	Interface usability (5)	Simple interface usability issues not related to broader functionality issues
Analytic Method	Am I doing it right? (8)	Concerns about doing the TSM method correctly
	Misunderstanding directions (34)	Disagreement, difference of opinion, or alternative interpretation of how the coding process was to be executed
	Comparison with other Instruments (19)	Comparing the TSM process with other teacher assessment systems
	Time (19)	How long or quickly the TSM process takes
	Perception (7)	Personal interpretation of data
	Cognitive load (6)	The mental capacity and attention required to complete the task
	Personal description of process (16)	Justification for a particular rating or lack thereof
	Interference (4)	Prior experience(s), tools, or methods, influencing one's rating
	Hard to find examples (19)	Difficult to find specific examples bound by time and space of attributes
	Forcing (7)	Tendency to force the TSM model onto the videos
	Concerns (26)	Concerns about how the TSM method would be used and applied
	Challenges of outside evaluator (18)	The difficulty of having an outside evaluator use this TSM method

Table 2, continued.

Coding Categories, Definitions, and Frequencies.

Category (Freq.)	Codes (Freq.)	Definitions
	Skill (13)	The level of expertise required to rate/ analyze a certain piece of evidence
	Continuum (24)	Issues related to putting practices on a continuum of expertise
	Benefits (22)	Benefits from using this TSM method, including benefits to the school or individuals
	Collaboration (9)	The effect on a teacher's evaluation of a video when collaborating with others
	Evidence (120)	Statements about what kinds of evidence best describe particular attributes
TSM Lens	Overlapping domains (22)	Attributes and domains that overlap in Practice
	Outside domains (2)	Finding evidence of "good teaching" that did not fit any of the lens attributes
	Wording (23)	Particular wording that was liked or not liked

Figure Caption

Figure 1. An example *Content & Curriculum* attribute in the Teacher Success Model (TSM)

with the accompanying indicators for each level of the continuum of teacher practice.

Figure 2. The Video Analysis Tool.

1 Domain 1: Content & Curriculum: Teachers demonstrate a strong knowledge of content area(s) appropriate for their certification levels.

2

Level	Basic	Proficient	Advanced
3 **Attribute A.** Demonstrates content knowledge in teaching area	**5** Teacher utilizes current text to develop teaching based lessons for students. Teacher is able to identify state standards included in lessons. Teaching is directly related to the current content for the grade level and content area. Teacher maintains a consistent focus on the content contained in the school syllabus.	Teacher supplements current text based on understanding of the content needed for overall class success. Teacher prioritizes state standards to focus on the content needed most by students. Teacher seeks deeper knowledge of content to improve student success. Teacher is open to suggestions of how the local curriculum can be adapted to include additional opportunities for students.	Teacher adapts content to address individual student needs. Teacher adapts content dynamically to address individual student needs. Teacher develops advanced, in-depth content knowledge Teacher develops novel approaches based on emergent opportunities.

1 = GSTEP Domain (1 of 6); 2 = Continuum of Teacher Practice; 3 = Teaching Attribute; 4 = Indicators of evidence for attribute

a.) Video player (view)

b.) Comments pane (annotate)

c.) TSM lens (assess)