The Role of Video in Early Field Experiences

Laura H. Baecher
CUNY Hunter College

Shiao-Chuan Kung
CUNY Hunter College

Abigail M. Jewkes
East Carolina University

Christine Rosalia
CUNY Hunter College

Recommended Citation
The Role of Video for Self-Evaluation in Early Field Experiences

1. Introduction

University-based teacher education programs are increasingly concerned about designing meaningful early field experiences (EFEs; Darling-Hammond, 2010). Well-designed EFEs provide targeted feedback on candidates’ performance using evaluation criteria aligned with those that will be encountered in subsequent teaching (Grossman, 2010; Poulou, 2007). By incorporating self, as well as supervisor, evaluation early in their professional preparation programs, candidates gain practice in applying standards to their teaching that encourages an active stance in the evaluation process. Self-evaluation is a well-established means for students to recognize their strengths and weaknesses, increase personal responsibility for their learning, and leads to more focused goal-setting (Andrade & Valtcheva, 2009; Price, 2005). Self-evaluation is widely considered essential to long-term professional growth (Taras, 2010), but without clear expectations aligned to performance, it may be inaccurate. The use of rubrics and video models, as well as video records of one’s teaching, can help strengthen the accuracy of the self-evaluation process (Kong, Shroff, & Hung, 2009).

In this study, we developed a conceptual framework outlining how video may influence the self-evaluation process in a way that positively supports teacher candidates in the early stages of their development. The research described in this paper grew from a recognition that although candidates in our teacher preparation program were being evaluated on their teaching performance by university supervisors during EFEs, self-evaluation was largely incidental. We therefore experimented with requiring self-assessment as part of the EFE evaluation of teaching to determine its efficacy.
In our teacher education program, first-semester teacher candidates plan and video record themselves teaching a 20-minute mini-lesson to a small group of learners at a local PreK-12 school as part of an EFE. This video is then submitted to a supervisor who evaluates performance according to an institutionally-developed observation rubric. For this study, prior to this EFE evaluation, we introduced two classes of candidates to samples of previous semester teacher candidates’ supervisor-completed evaluation rubrics that were given for their EFE teaching. One class received videos of teaching that corresponded to those evaluation rubrics while the other class was given descriptive narratives of the lessons that corresponded to the rubrics. This approach was designed to support all teacher candidates’ readiness for the video-based evaluation process within the EFE in order to enhance their ability to self-evaluate realistically, while at the same time experimenting with the role of video models in that process.

Video records of teaching have been used successfully to support reflection in pre-service teacher development programs worldwide (Eröz-Tuğ, 2013; Harford & MacRuairc, 2008; Kong, 2010; Masats & Dooly, 2011; Rosaen, Lundeberg, Cooper, Fritzen & Terpstra, 2008; Sonmez & Hakverdi-Can, 2012; Welsch & Devlin, 2007). For example, a study involving peer video-recording and group discussion of teacher candidates’ classroom performance in local schools conducted in Ireland showed that these activities exposed pre-service teachers to a range of diverse teaching methodologies and fostered self-reflection (Harford & MacRuairc, 2008). In a related study by Kong (2010), teacher candidates in Hong Kong captured videos of their practice and used these video records to conduct self-reflections of their curriculum planning and evaluation, student interactions, discipline and classroom management. In this study, a rubric containing specific items to observe in each dimension was provided to guide students in their reflections. The researcher found that teacher candidates generated more and deeper reflective
notes after watching clips of their own teaching when using this protocol. Additionally, in Turkey, Sonmez and Hakverdi-Can (2012) found that video models heightened teacher candidates’ ability to observe and evaluate teaching practice. However, while the combination of rubrics and video for self-evaluation has been previously studied, its use in preparing teacher candidates for early fieldwork experiences has been less explored. The match between evaluation of performance on video between teacher candidates (novice observers) compared to the observations of their supervisors (more experienced observers) is also of value to the implementation of video for self-evaluation. Thus, our inquiry sought to broaden the literature in these areas and was guided by the questions below:

(1) How did the review of rubric-anchored video models influence teacher candidates’ capacity to self-evaluate?

(2) How did the process of reviewing video models influence teacher candidates’ thinking about their own teaching performance?

2. Conceptual Framework

The process of using video both as a model and a vehicle to support teacher candidates in self-evaluation depends on a number of interrelated steps that can be facilitated by teacher educators. In designing the present study, we developed a conceptual framework outlining how video might influence the self-evaluation process in EFEs (see Figure 1). Drawing upon a number of theories about self-assessment, the use of video for self-review, and the practicalities of the need for training in the use and application of observation rubrics, this framework attempts to make explicit connections between theory and practice, as further explicated below.
2.1 Invoking social comparison via video models.

The use of video cases, in addition to written materials, is known to be a powerful instructional tool for teacher educators (Boling, 2007; Koc, Peker, & Osmanoglu, 2009; Masats & Dooly, 2011; Wang & Hartley, 2003; Welsch & Devlin, 2007). Intense viewing of video cases helps novices examine classroom incidents (Sherin, 2000), provides a shared common experience in a controlled environment (Sonmez & Hakverdi, 2012), and aids in establishing a view of teaching as the complex interaction of a number of factors. Masingila and Doerr (2002) reported that multi-media cases guided teacher candidates’ instructional practices and helped them connect their own practice with that of the teacher in the video case.

Reviewing models of desired performance, such as are found in video cases, can also decrease the likelihood of leniency in self-evaluation (Heidemeier & Moser, 2009). Social comparison theory (Festinger, 1954) suggests that we seek to understand the level of our own performance by using the performance of others as a point of reference. When discrepancies arise, for instance when we notice that others appear to be performing at a higher level, this triggers a desire to examine our own behavior (Suls, Martin, & Wheeler, 2002). Studies show that we are inspired and motivated to monitor our own behavior and make adjustments when presented with models of performance that demonstrate an achievable level of performance. On the other hand, “superstar” models of performance appear to deflate and demoralize (Lockwood & Kunda, 1997, p. 101). Thus, an important consideration when developing video models of teaching is to ensure they are similar enough to teacher candidates’ abilities, while also introducing new ideas to help them learn and grow.

For these reasons, the video models utilized in our study were from former teacher candidates from the same program, recorded when they had been at the same point in their
graduate careers as the study participants and in the same partner school fieldwork sites. This provided current teacher candidates a glimpse of attainable performance achieved by peers, thus invoking the benefits of social comparison as a motivator for performance. Additionally, research suggests that the more familiar the teaching context is with the viewers’ own classroom environments, the more open they will be to learning and applying what they see (Sherin & van Es, 2005). We hoped that by providing video models of peers in similar fieldwork settings, delivering lessons that were evaluated as meeting standard, we could capitalize on the tendency to draw parallels between one’s own performance and that of others, thus energizing the self-evaluation process.

2.2 Generating cognitive dissonance via supervisor-completed rubrics with video records.

One challenge of using self-evaluation with teacher candidates is the possibility that they will perceive their performance in a more favorable light than their supervisors. Sitzmann, Ely, Brown and Bauer (2010) state:

Accurately self-assessing one’s performance is intrinsically difficult. Learners must have a clear understanding of the performance domain in question. They also must be willing to consider all aspects of their knowledge levels (not just the favorable) and to overcome the egocentrism that results in people assuming they are above average in most aspects of their lives (p. 180).

Over- and under-rating in self-evaluation arises from several sources. For instance, some self-evaluators show a leniency bias, displaying the tendency to rate oneself more favorably than others might (Falchikov & Boud, 1989; Seidel, Stürmer, Blomberg, Kobarg, & Schwindt, 2010), while others display a modesty bias, in which performance is underrated (Farh, Dobbins, & Cheng, 1991). These biases tend to occur more in situations where the self-evaluation is for
evaluation and promotion purposes, rather than simply for developmental purposes (Zimmerman, Mount, & Goff, 2008).

One way to increase the validity of self-evaluation (and to lessen bias) is to provide raters with a clear description of performance criteria. If a gap appears between what a supervisor sees and what a teacher candidate sees, a state of cognitive disequilibrium (Piaget, 1929) or dissonance (Festinger, 1957) may occur. Psychologists have theorized that because humans seek to maintain consonance between their actions and beliefs, dissonance is induced when one is presented with an inconsistency between an action taken and one’s beliefs about that action. The contradictory situation produces feelings of discomfort, and this theory posits that an individual will change beliefs or behaviors in order to reclaim a sense of internal consistency. Cognitive dissonance can be expected when teachers need to reconcile a supervisor’s evaluation that is different from their own. In children’s struggle to make sense of new and contradictory information Piaget saw the fuel for learning and development, while Festinger’s notion of cognitive dissonance implies the same for adult learners.

In a study of self-evaluation combined with supervisor evaluation, Ross and Bruce (2007) found that supervisors contributed to teacher self-assessment by clarifying goals (i.e. dimensions of teaching that define excellence) and criteria (levels of performance that constitute a hierarchy of professional growth). They also provided credible feedback on whether particular standards of teaching have been achieved (e.g. through classroom observation). Essentially, when supervisors provide evaluations that are compared to self-evaluations, supervisees become clearer about the performance criteria (Conway & Huffcutt, 1997). Clearly understood rubrics “give students valuable information about the task they are about to undertake and take the
guess-work out of understanding their learning targets, or what counts as high quality work” (Andrade & Valtcheva, 2009, p. 13).

Based on the research highlighting the need to compare self to supervisor ratings, in our study we asked teacher candidates to view sample videos and supervisor evaluations and then reflect on any discrepancies between their impressions of the lesson and the supervisors’. We hypothesized that teacher candidates would need to adjust their beliefs about the teaching they viewed in relation to the experts’ scores. The focus on refining how one rates someone else’s video, we hoped, would translate to less bias when teacher candidates assessed their own teaching videos.

2.3 Self-evaluating teaching using one’s own video records.

While teachers and teacher candidates regularly confront high-stakes evaluation based on classroom performance, they unfortunately also experience evaluation as political enactment of “control-oriented routines” (Towndrow & Tan, 2009, p. 285) rather than a vehicle for professional development. Therefore, inviting teacher candidates to be evaluators of their own performance could encourage teachers to be active agents in the evaluation process (Cheung, 2009; Cranton, 2001; Kong, 2010). By harnessing the power of review to “notice, revisit, and investigate” (Rosaen et al., 2008, p. 356), teachers arrive at a more fine-grained understanding of their performance than memory-based recall. Furthermore, when teachers have an opportunity to watch and analyze video of their own rather than others teaching, they may experience greater motivation and engagement in the activity, further enriching the reflective process (Seidel et al, 2011).

Support for self-evaluation in learning from video or other performance artifacts is based on theories of self-regulation as connected to metacognition (Zimmerman, 2008). Because
engaging candidates in the metacognitive task of stepping back to appraise their own performance stimulates self-monitoring (Bransford, Brown, & Cocking, 1999; Winne & Hadwin, 1998), self-evaluation encourages a more active and self-reliant role in one’s own learning (Wiggins, 1998). Heightened self-awareness, leading to ever greater self-regulation, has been shown to yield benefits including a greater awareness of the evaluation process and scoring criteria, an increased understanding of the instructional content, and improved performance (Davies, 2002; Fallows & Chandramohan, 2001).

To apply the literature on video as a tool in self-evaluation to our study, teacher candidates evaluated their teaching performance via video using both the rubric and open-ended reflective writing. In our conceptual framework, it was anticipated that this stage of self-evaluation would yield more realistic views on performance. In sum, by self-evaluating after having seen video models and adjusting their perceptions to that of the supervisors’ scores, we hypothesized that the teacher candidate would self-assess their own videos of teaching with less personal bias and with more responsiveness to understood criteria.

5. Methods

5.1 Overview of the design of the study

In order to investigate the impact of video models on teacher candidates’ readiness for and capacity to self-evaluate their teaching performance in an EFE lesson, we employed a mixed methods design. Our study centered on teacher candidates (N=31) from two sections of a first semester TESOL (Teaching English to Speakers of Other Languages) methods course. All teacher candidates were required to spend 35 hours as a classroom assistant in a P-12 setting, where they provided tutoring and small-group instruction and were asked to submit one 20-minute video of themselves teaching a mini-lesson.
First, one class was required to read descriptions of three lessons, from three different teacher candidates, which had each been rated as meeting standard in relation to this EFE observation rubric. The other class section, instead of reading written descriptions, viewed online videos that corresponded to these three lessons. Both sections therefore received preparation materials for their video-of-self task, but one had access to video models in addition to written materials. Once data was collected for each group, the teacher candidates who did not first view the video models were then allowed to do so. Participants in both groups wrote reflections on their impressions of the lessons and their views on how the supervisor had evaluated the lessons.

Next, all teacher candidates selected a video-recorded lesson from their EFE site. Teacher candidates were encouraged to record more than one sample video, and thus, most candidates were self-selecting their 20-minute lesson from a personal library of their teaching. Candidates wrote reflections on their own teaching as viewed in their videos and on the process of video analysis of their own teaching.

Finally, the chosen video lessons were first self-rated by the teacher candidate and then rated by a university field supervisor according to an EFE observation rubric (see Appendix A). These scores were later analyzed statistically to investigate which of the two treatments (videos or text descriptions) led to more accurate self-evaluation. Figure 2 graphically represents the procedures involved in the implementation of the design.

5.2 Participants

Participants in the study were teacher candidates at a large, urban college of education enrolled in a master’s degree program leading to P-12 state teaching certification in TESOL. All
candidates were enrolled in one of two sections of a required TESOL methods course during the same semester. As a required early fieldwork component of this course, these candidates were placed in the local public schools and served as classroom teaching assistants one to two days per week for a 15-week semester in an elementary or secondary classroom, under the guidance of a cooperating teacher. The video-model course section had 16 participants and the text-only section had 15 participants. Teacher candidates in both sections of the course were invited to participate; none were excluded based on performance in coursework or for any other reasons. In addition, formal ethics board approval was obtained, and consent for participation in this research study was not sought until course grades were submitted.

Participants ranged from 24-45 years of age, with zero to three prior years of teaching experience. Six participants held a prior teaching certification. The video-model section of the course was led by the first author, who had taught the course for five years; the other instructor had taught the course twice before. Both instructors also served as field supervisors for their class of teacher candidates. Both had prior experience evaluating teacher candidates in EFEs, and both participated in the development of the syllabus, which was identical for both classes.

5.3 Data Sources

Materials that were collected and analyzed for this study consisted of:

1. Teacher candidates’ written reflections in response to either viewing videos of three EFE lessons and reading accompanying evaluation rubrics (video-model class) or reading descriptions of the lessons and rubric evaluations of these three EFE lessons (text-only class). The three videos were selected by the methods instructors from video clips that had been submitted for this same assignment in prior semesters. University supervisors rated the selected videos as meeting standard, a 3 on a 1-4 scale. Both classes of teacher candidates wrote a 1-2
page reflection focusing on TESOL-specific considerations for their own lesson planning and how reading the rubric evaluations assisted them in doing so. The video-model class reflected on the benefits of watching the video lessons/cases, in addition to reading the evaluations.

2. Teacher candidates’ written reflections in response to viewing their own teaching on video. Both classes of teacher candidates wrote a 1-2 page reflection focusing on their language and content learning objectives, lessons learned, and their experience self-reviewing their teaching on video.

3. Teacher candidates’ self-ratings using the EFE observation rubric. This rubric was developed to meet two needs. First, there was a need to provide performance feedback for the observed lesson taught in the EFE because this had not been a formalized observation up to that point. Second, the EFE rubric was a scaffold to a much longer (50-item) rubric that teacher candidates would need to negotiate in their final semesters in the program. This EFE rubric served as a scaffold by using the same domains as the student teaching observation rubric, but reducing the items to account for the modified expectations for teaching. Faculty developed both the EFE and the student teaching rubrics, and the EFE rubric had been piloted and revised over four semesters prior to this study. The EFE observation rubric consisted of 20 statements, which were each rated using a 4-point scale: (1) indicator not met; (2) indicator partially met; (3) indicator met; or (4) indicator exceeded. A zero option, which was not counted in scoring, was provided for instances where respondents did not feel there was sufficient evidence to make a judgment. The four points of the scale align with the scale used on the subsequent, more extensive observation rubric used in student teaching, as well as with the scale used in an observation framework that teacher candidates encounter in the local school system (Danielson, 2013),
4. Supervisors’ scores of candidates’ performance on the EFE observation rubric based on the video-recorded lesson (experimental and comparison classes). Supervisors and candidates scored their videos without having seen one another’s ratings. The supervisors, who were the instructors of the classes, scored the teacher candidates separately, and then these scores were averaged for a single supervisor rating on each item, on a scale from 0-4.

5.4 Data analysis

Teacher candidate and supervisor ratings on the EFE observation rubric were entered and analyzed using SPSS. Given our small sample size, correlations and the Wilcoxon Signed Rank Test (a non-parametric test) were conducted to compare (a) teacher candidate and supervisor ratings and (b) video-model and text-only classes.

Teacher candidates’ written reflections (data sources 1 and 2) were consensus coded by the first two authors. First, these were open-coded by looking for emerging themes while reading responses to each of the questions (Corbin & Strauss, 2008). Many of the codes were in vivo codes that “capture the actual words used by participants” (Leech & Onwuegbuzie, 2007, p. 576). Finally, the categories were selectively coded (Creswell & Clark, 2007) by rereading through all of the data to make sure the previous codes and categories were analyzed according to similarities and differences among participants and connected to emerging theory based on the principles of grounded theory (Glaser & Strauss, 1967). When appropriate, original text found in the data was excerpted to match each theme. An example theme was “noticed discrepancy between expert and own assessment.” Its matching instance was “I was surprised to see that she received ‘indicator not met’ on the rubric.” Inadequate themes were discarded or replaced, and discrepancies were discussed until all segments of the written reflections were coded.

6. Findings
6.1 Teacher candidate and supervisor ratings on EFE observation rubric

Teacher candidates in both classes rated themselves numerically higher than did their supervisors, consistent with the leniency trend noted in the literature on self-evaluation (see Table 1). Variability in ratings between the two teacher candidate groups was very small (.35 and .32), suggesting similarities in performance among candidates in both classes. However, teacher candidates in the video-model class rated themselves slightly lower than did those in the comparison group (2.66 versus 3.27 on a 4-point scale). The comparison class rated themselves higher than the video-model class did, and higher than their supervisors did. In addition, full sample analysis of teacher candidates’ rubric ratings compared with their supervisors’ ratings indicated a moderately significant correlation \(r = 0.37, p = 0.042\) indicating that, in this study, supervisors and teacher candidates were fairly similar in their ratings.

In contrast to the teacher candidates’ scores, supervisors’ mean scores were similar across both groups and exhibited less variability than teacher candidates’ (see Table 1). Correlations between the teacher candidates’ and supervisors’ scores in the video-model group were moderate \(r = 0.33\), while they were small \(r = 0.11\) in the comparison group; neither was statistically significant (although this finding could be due in part to the small sample size). Comparisons using the Wilcoxon Signed Rank Test revealed a significant difference in the teacher candidates’ and the supervisors’ scores in the comparison group \([Z(15) = -3.295, p = 0.001]\), but not for the video-model group \([Z(15) = -0.45, p = 0.649]\). Teacher candidates who viewed video cases prior to their videotaped teaching scored themselves more similarly to their supervisors than the teacher candidates who read descriptions of lessons and reviewed rubrics without watching the videos.

[Insert Table 1 about here]
6.2 Preparing for the EFE teaching performance task

Candidates’ written reflections about the video models and rubrics, which were ungraded, informal writing submissions, revealed the value of this activity. Candidates expressed their belief that the opportunity to read and review evaluations of teaching from the same EFE teaching performance task they were about to begin was highly valuable. Both groups of teacher candidates stated that this gave them a chance to understand what was expected, what the rubric looked like, and what the supervisor was likely to focus on. Overall both classes were very positive about how this prepared them for their own teaching task.

One theme that emerged from reflections written by the teacher candidates in the experimental class was that reviewing the video models along with their EFE observation rubrics eased them into the process of preparing for their own video observation. One participant stated, “I’m confident that the result of my presentation was better than what it would have been without seeing these videos... Overall, the videos were very helpful.” Another teacher candidate wrote, “I’ve found watching other videos has been extremely useful to becoming aware of my own development as a teacher.” Model videos were reported to be “useful” in several ways. First, they were helpful when teacher candidates were planning their own lessons. Teacher candidates made adjustments to what to focus on and what problem areas to avoid. In her reflection, one teacher candidate wrote:

As I was creating my lesson, I tried to incorporate certain components that I thought were important as well as to avoid any problem areas that I saw in the videos… After the observations, I made sure to incorporate key components into my lesson, such as adapting content to provide comprehensible input, building schema, creating a hook for the beginning of the lesson.
The videos also helped teacher candidates better understand what level of performance was expected of them and how the observation rubric worked. One candidate stated, “While I had a pretty good idea of what I wanted my lesson to look like, watching the videos of past students’ mini-lessons provided a lot of insight that allowed me to alter aspects of my lesson to better suit the requirements of this assignment.” Another indicated that, “Watching the videos and reading the attached rubrics helped me to understand how my lesson plan would be evaluated.” Some teacher candidates also found the videos to be motivational by stating, “It was refreshing to see these other students in the midst of their video of teaching assignment and it encouraged me to go bigger, aim higher, and achieve what I saw as further possibilities in these videos.” Others made connections between the content of the videos and concepts and techniques they were learning in their courses as indicated by the comment “I felt that she took a lot of time just getting kids settled and quiet. Maybe some of the techniques we talked about in class (clapping your hands, for example) would have been helpful to implement here.”

Another main theme that surfaced from the reflections is the discrepancy in what some teacher candidates saw in the videos and the accompanying rubric scores. One teacher candidate said, “There were three areas that I thought were very successful that the rubric indicated as ‘not met’.” This discovery caused both surprise and anxiety. “I was really surprised at first to see the low grade on the rubric. It actually made me pretty nervous because I thought the lesson had gone well,” expressed one candidate. These comments revealed that the videos generated a sense of cognitive dissonance—between how they viewed the teaching and how the supervisor viewed it—that spurred candidates to review the criteria for performance in relation to the lesson viewed. They did not want a gap between their ratings and a supervisor’s to persist into their own later teaching, and perhaps were more ready to address and reconcile the dissonance they
experienced since the teacher being evaluated was another and not themselves. Rather than being defensive of the teacher, they willingly adjusted their perceptions to align more closely to the supervisor.

In a similar way to the video model class described above, the teacher candidates who read the evaluations and class descriptions rather than viewing videos also found the exercise worthwhile. They became aware of areas to pay attention to and mistakes to avoid. They also felt that having access to the rubrics helped them gain understanding of how they would be evaluated. One candidate stated, “After reading the reports I have understood what skills I need to focus on while getting prepared for the task. Now I know what is expected of me and I will try to meet the expectation.” However, among these candidates, reading the supervisor-completed evaluations did not appear to excite any kind of emotional response. These teacher candidates, in contrast to those in the video-model class, did not express motivation or anxiety after reading about others who had gone through the experience for which they were preparing. In fact, it appeared that the teacher candidates in this group used the materials they were given simply as data to be analyzed. They noticed the indicators least frequently met by the student teachers they read about and used that information to prepare for their own lessons. For example, one of the teacher candidates in the comparison class said:

By comparing the three charts, I quickly found that there are three indicators that all three teachers failed to meet… These are the three areas that I will be sure to pay attention to in my lesson… I will try to use backwards design with clear objectives in mind when I plan my lesson… I'll make sure to find out their names, where they’re from, and start things off in a friendly manner, to make them feel comfortable.
In summary, teacher candidates’ reflections on the models they reviewed demonstrated that both groups benefited from watching other teacher candidates’ videos and/or reading the written descriptions and evaluations. However, there were two important differences in the reflections: (1) the teacher candidates who viewed the video models became aware of the discrepancy between their own assessment using the evaluation rubric and a supervisor’s ratings, and (2) these teacher candidates also expressed more anxiety and excitement about their upcoming observation than the teacher candidates who read descriptions and rubric evaluations. This may have been due to a greater intensity to the experience of seeing the teacher in the video lesson (social comparison) and greater discord at discovering differences in the supervisor’s evaluation and their own (cognitive dissonance) after viewing the video and formulating their own opinion. Perhaps reading the description of the lesson and its corresponding evaluation created cognitive dissonance, but there was less indication of the dissonance being transformative. Teacher candidates did not appear to make a personal connection to the teacher in the description, rather they focused on how they would be different—perhaps avoiding the possibility that they could inadvertently make the same teaching “mistakes”.

6.3 Participation in video analysis of self across experimental and comparison classes

After both groups of teacher candidates taught the lessons that were video-recorded and evaluated by themselves and their supervisors, they reflected on their teaching performance. In the reflection papers the experimental class wrote, candidates made frequent references to their performance in relation to the categories of the observation rubric. For example, one participant in the experimental class wrote:

The part of the lesson where I would check for comprehension after demonstrating the poster with the target language that I created was missing. The students did get a chance
to practice the target form when they were writing their own directions in groups, which I observed while I monitored the groups and helped them with the assignment (indicators 4a and 4c). I still modeled to the students what they were expected to perform, and I tried to lead an engaging lesson. I also tried to be enthusiastic and inspiring, and the students participated in the lesson (indicators 4a and 7b).

One theme that emerged from the self-reflections of teacher candidates who had seen the model videos was the value of watching videos of one’s teaching performance as it aided in shifting the focus away from the teacher and towards the learners. For example, one participant stated:

After watching myself teach, I always walk away finding areas to improve on and ideas on how to teach the lesson differently the next time around… While you are teaching, even though you are learning what may or may not work for the next time you teach, there is a lot going on in your mind and you are not able to step out of the circumstances at hand. When I watched the video I was able to take the lesson for what it was. I was able to see if the information was presented in a coherent manner and also if the class was responsive or not.

Candidates in the video-model class provided future-oriented comments, often referencing what they would do differently in their written reflections on their teaching video. For instance, one candidate commented: “In the future I will model the target language objective throughout the whole lesson… In the future I am going to be very aware of the visual cues I am giving and make sure they cannot be interpreted for something else.”

After viewing their lessons on video, candidates in the comparison class also reflected on aspects of their teaching in relation to indicators of the rubric. However, few teacher candidates
in the comparison class made plans for future lessons and those who did, made plans that were less specific than plans made by the experimental group. For example, they said, “In the future I would try to talk a bit less” and “Another part of the lesson that I would do differently next time involves the difficulty of the writing sample.” This seemed to be related to candidates in the comparison class making more generalizations about their effectiveness: “I thought the lesson was pulled off successfully” and “The students were very engaged in the lesson… All the students achieved the goal I set for them.” Such overtly confident assertions were fewer in the experimental group’s reflections. One more important difference between the two groups is that among the comparison group candidates, no one commented on the value of video as a tool for self-reflection and professional growth.

7.0 Discussion

We initiated our inquiry by introducing video models to one group of first-semester teacher candidates with the belief that this might positively impact preparation for their EFE teaching observation, as well as foster a greater ability to apply the evaluation criteria to their own performance. Our findings suggest that teacher candidates benefited by viewing model videos in preparing for their teaching experience. When it was time to assess their own teaching, they felt more prepared to do so and rated themselves more closely to their supervisors than did the group did who had access to written descriptions rather than video models. In addition, candidates who had access to the video models and rubrics did not overestimate their performance. Another benefit of viewing sample videos was how this appeared to encourage teacher candidates to recognize the value of recording and watching videos of teaching in their professional development. These results are especially notable given the study’s small sample
size, though it should be noted that there may be other influences on candidates’ ratings that were not accounted for in our study.

While both groups of teacher candidates reported that access to supervisors’ completed observation rubrics aided them in planning their lessons, the group that viewed video models seemed able to later apply this knowledge when they engaged in self-evaluation. Seeing the teaching enacted seemed to add a meaningful component to their interpretation of the supervisors’ ratings, which is consistent with the research on the need for scaffolding in preparing teachers to self-evaluate (Airasian & Gullickson, 1997; Ross & Bruce, 2007). The two-dimensional written descriptions of the lessons and supervisors’ ratings were enhanced with the additional dimension of video, which has been noted in research on the effectiveness of video cases (Boling, 2007).

Returning to our conceptual framework and beginning with Step 1 (see Figure 1), videos from previous teacher candidates who were successful at the task were viewed by the experimental group. Social comparison may also have been more fully invoked in the experimental group as the viewing of teacher candidates from the same program in the videos provided an accessible model for the viewers. Without a model to visualize in relation to their own performance, teacher candidates who did not have access to the video cases may have been less able to understand what teacher behaviors would be notable on the part of supervisors. By providing, as in step 2 of the framework, video-anchored models of what meeting performance looked like across different teachers, the rubric was made more comprehensible to teacher candidates.

Recent research on interrater reliability in video-based evaluations of teaching indicates that a good deal of training using expert score comparisons is necessary, even among highly
experienced classroom observers (Hill, Charalambous & Kraft, 2012). This suggests that in our study, viewing the video cases may have had a norming effect on this class, as these candidates adjusted their prior beliefs about the quality of teaching in the videos to conform with the rubric scores they were provided. Only reading about, but not viewing, “meeting standard” did not provide teacher candidates an opportunity to hold their own image of meeting standard against a benchmark. Especially notable is the cognitive dissonance experienced by teacher candidates who were able to construct their own beliefs about how the model lessons might be scored prior to weighing those against the supervisors’ ratings. In contrast, the comparison group teacher candidates were only able to interpret the events as the supervisor had seen them. The teacher candidates who only read the written descriptions of the lessons did not appear to anticipate that their assessment might differ from the assessment of experts; hence no dissonance, or subsequent self-questioning, appeared to be experienced. The opportunity that video-review provides to generate dissonance has been a notable feature in its high degree of impact on teacher candidate reflection (Rosaen et al., 2008).

In the third step of this project, teacher candidates self-evaluated their teaching performance prior to receiving feedback from their supervisor, using both the rubric and open-ended reflective writing. The data from this phase indicated that the teacher candidates who had access to the video cases were able to align their self-evaluations more with their supervisor’s, and perhaps to be more cautious about jumping to conclusions about the quality of student learning. In a sense, the whole model is predicated upon the use of video to generate greater reserve in self-rating, when viewing the teaching of oneself or others. Such experiences may influence candidates’ stance towards supervisory feedback and enhance a critical approach to their own teaching.
8.0 Conclusion

Across institutions of teacher preparation, the ease, affordability, and accessibility of digital video records of teaching will continue to make video an increasingly essential educational tool. This study suggests that video teaching models accompanied with rubric evaluations introduced early in the teacher education program is one way to develop teacher candidates’ emergent understanding of how classroom practice will be evaluated. Equally important is the opportunity for teacher candidates to discuss and reflect on these models, especially in relation to their videotaped practice. Rather than leaving candidates unsure about expectations for teaching in early fieldwork, focused video analysis is a means to innovate course design and promote teachers’ self-awareness of their practice. Providing opportunities for candidates and faculty to view classroom teaching videos from their local context, then rate independently and discuss these ratings serves to demystify the evaluative process. This work could also build readiness for self-evaluation in the teacher education program, which is important in all contexts for teacher preparation.

For EFE supervisors, incorporating scaffolded opportunities for candidates to view and reflect on their teaching may result in stronger self-evaluation capabilities when candidates reach their practicum year. This study suggests that the opportunity to view videos of competent, proximal others in similar teaching contexts, alongside reviews of these teaching events generated by more experienced others, may create positive social motivation and clarify performance expectations. In combination, such activities may drive down inflated self-ratings and open up new ways for teachers to evaluate their own performance. Our results indicated that teacher candidates not only appreciated pre-rated model cases of teachers who were ahead of them in their coursework, but that they were more introspective of what the assessment tool and
use of video could offer in their continued teacher development. The use of peer evaluation could also enhance and scaffold the path from supervisor to self-evaluation.

This research takes a step towards understanding what role video might play in early field experience teaching. Further investigation is needed into the processes invoked when video is introduced into preparation activities, how it may influence self-evaluation, and in what ways its use could create change in the traditional supervisor-candidate feedback process. As teachers’ performance is regularly assessed by others, equipping candidates with the tools for meaningful self-evaluation may support their empowerment in those processes. If teacher candidates create video records of their teaching, and are able to realistically evaluate their own performance, supervisors may be able to be utilized to a greater extent in a mentoring or coaching role. Snyder (2009) suggests that an overemphasis on assessment of performance, at the expense of development, may ultimately shortchange teacher candidates just beginning to develop the habits of mind of reflective practitioners. Candidates’ ability to self-evaluate in ways more similar to their supervisors might open up new possibilities for the role of video in the supervisory feedback process.
References


Table 1

Comparison of Self and Supervisor Rubric Ratings across Classes

<table>
<thead>
<tr>
<th>Classes</th>
<th>Teacher Candidates’ self-ratings</th>
<th>Supervisors’ ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n$</td>
<td>$M$</td>
</tr>
<tr>
<td>Experimental class</td>
<td>16</td>
<td>2.66</td>
</tr>
<tr>
<td>Comparison class</td>
<td>15</td>
<td>3.26</td>
</tr>
</tbody>
</table>
Figure 1. Conceptual framework of video-informed self-evaluation
Figure 2. Research process flow chart
## Appendix A
### EFE Observation of Teaching Rubric

Instructions: Based on your observation of this teacher candidate’s pre-student teaching performance, please assess his/her development of the following competencies by completing this form.

<table>
<thead>
<tr>
<th>Performance Indicators</th>
<th>Indicator not met: Performance is less than adequate; considerable improvement needs to be made in this area. (1 point)</th>
<th>Indicator partially met: Performance is not consistent; periodic checks will be needed during student teaching (2 points)</th>
<th>Indicator met: Skill performed at appropriate level; meets expectations at this point (3 points)</th>
<th>Indicator exceeded: Performance is superior; exceeds expectations at this point (4 points)</th>
<th>No Opportunity to Evaluate</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Knowledge of Learners</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1a. Adapted discourse to provide comprehensible input appropriate to language proficiency and age level of learners</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>1b. Made links to students’ background, culture, or native language where appropriate</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>1c. Called on students by name</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td><strong>2. Knowledge and Use of Subject Matter</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2a. Selected content appropriate to developmental and language proficiency level</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>2b. Introduced content by activating/building schema</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td><strong>3. Skill in Planning</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3a. Developed clear, appropriate and meaningful content objective, worded as desired student understanding</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>3b. Developed clear, appropriate language objective, in the area of L, S, R, W, Grammar, Vocab, Pragmatics, or Language Lrng. strategy</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>3c. Selected/created appropriate and engaging instructional resources, visuals, materials, and/or technologies</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>3d. Adapted/modified text or materials to make the content accessible to ELLs</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td><strong>4. Skill in Teaching</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4a. Used a “hook” to interest students in the content &amp; motivate them</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>4b. Modeled task before asking students to perform</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
<tr>
<td>4c. Teacher encouraged wide number of student responders, through use of built-in wait time (think/pair/share, turn and talk)</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
</tbody>
</table>
4. Pupils engaged in using/practicing the target language structure or feature during most of the instructional time

5. Skill in Developing Caring Learning Environments and Showing Respect for Learners
   5a. Showed rapport with students
   5b. Listened empathetically and thoughtfully responded to student contributions
   5c. Used positive reinforcement to focus students on learning

6. Skill in Assessment
   6a. Checked to see if students understood information, especially by circulating room or noticing students performance
   6b. Provided helpful feedback to students

7. Professionalism
   7a. Communicated clearly and accurately, as a model of speaking, listening, reading and writing
   7b. Demonstrated enthusiasm and poise throughout the lesson