Inspiring Critical Reflection in Preservice Teachers

Susan E. King

Abstract

The purpose of this study was to determine the effectiveness of peer-generated feedback versus visual feedback in inspiring preservice physical education teachers to produce reflections that are more critical than technical. The participants were junior-, senior-, and graduate-level health and physical education majors enrolled in an instructional methods course. Upon completion of four required teaching labs, participants were given either peer-generated feedback or a visual record of their teaching. They received reflective prompts and were asked to reflect upon their teaching. Among participants in this study, 61% of more critical, intermediate reflections were based upon visual feedback. Peer-generated feedback produced 73% of technical, novice reflections. It was concluded that visual feedback inspired more critical than technical reflection.

According to educational researchers and teacher educators, both past and present, reflection is regarded as an essential teaching behavior (Crotty & Allyn, 2001; Dewey, 1933; Langley & Senne, 1997; Parkinson, 2005; Schon, 1983; Tsangaridou & O'Sullivan, 1997; van Manen, 1977; Zeichner, 1987). Reflection can be defined as active engagement of the mind to solve problems and improve performance in the classroom; a critical analysis involving self-understanding, heightened consciousness, and emancipatory learning (Dewey, 1933; van Manen, 1977; Zeichner & Liston, 1996). The historical roots of reflection lie in Dewey's (1933) view of reflective thinking as an essential tool in enabling educators to direct their activities and plan an effective course of action with appropriate ends in view. Reflection "converts action that is merely appetitive, blind, and impulsive into intelligent action" (Dewey, 1933, p. 17). Whereas non-reflective teaching can be described as unconscious, automatic, and narrow, practitioners who engage in reflection are those who do more than exhibit a set of teaching behaviors identified and prescribed by others (Dewey, 1933; Sparks-Langer & Colton, 1991; Zeichner & Liston, 1996). Instead they identify problems, question goals, explore contexts, analyze possibilities, and craft appropriate educational experiences to benefit learners.

Reflection in its lowest form focuses on the technical aspects of teaching, namely, the application of teaching skills in response to narrowly construed problems (Cruikshank, 1985; Hatton & Smith, 1994; Schon, 1983; van Manen, 1977; Zeichner & Liston, 1996). Technical reflections often employ "research-based means for achieving given ends" (Hatton & Smith, 1994, p. 3). The development of deeper forms of reflection has been a persistent challenge for teacher educators. Critical reflection is described as cognitive activity during which one carefully considers the impact of one's actions upon others, and provides a rationale, taking into account the social, cultural and/or political forces at work in the event (Hatton & Smith, 1994). Most preservice teachers exhibit technical rather than critical reflective skills and "many researchers concluded that developing deep levels of reflection is difficult for prospective teachers and generated few clues as to why this may be so and what can be done to better support their reflective development" (Risko et al., 2002, p. 136).

Although preservice teachers are slow to understand and value the benefits of reflection, development of reflective skills remains a key

component in most teacher education curricula. Many state licensing agencies and specialized professional associations, such as the National Association for Sport and Physical Education (NASPE), list reflection among competencies deemed essential for beginning teachers (National Association for Sport and Physical Education, 2001). By encouraging preservice teachers to engage in meaningful reflection, they can better assess and improve their pedagogy by linking theory to practice. According to Zeichner (1987), strategies to promote reflective abilities fall within six categories: 1) writings, 2) curriculum inquiry, 3) various supervisory approaches, 4) action research, 5) ethnography, and 6) reflective teaching. It appears that the most common strategies involve structured field experiences or microteaching accompanied by some form of writing (McCollum, 2002; Sparks-Langer & Colton, 1991; Tsangaridou & O'Sullivan, 1997; Tsangaridou & Siedentop, 1995). The result of these strategies is what Schon (1983) calls *reflection-on-action*, which takes place before or after the teaching act, as opposed to *reflection-in-action* which is an immediate response to classroom events.

As preservice teachers engage in structured field experiences, they generally receive feedback on their teaching. Prospective teachers must practice their skills "frequently under good conditions, and get help in the form of instruction, supervision, and feedback..." (Siedentop & Tannehill, 2000, p. 3). This feedback provides a stimulus for their reflection. Common forms of teaching feedback are written, verbal, and visual. Written and verbal feedback generally comes from a trusted "other," such as an instructor, cooperating teacher, or peer who serves as a coach (Hatton & Smith, 1994). After observation, coaches dialogue with practicing preservice teachers and provide comments verbally or in writing regarding their teaching performance. Preservice teachers then reflect upon the coaches' feedback. Visual feedback is a video recording of teaching. It enables practicing teachers to reflect on their performance privately or with help from a knowledgeable other. The use of written/verbal feedback and visual feedback are considered effective strategies to promote reflective activity (Crotty & Allyn, 2001; Hatton & Smith, 1994; Napper-Owen & McCallister, 2005; Schon, 1987). But one of these methods may be a more appropriate stimulus for reflection-on-action.

The purposes of this study were a) to examine the effects of peer-generated, written/verbal feedback as well as visual feedback to determine whether one type is a superior stimulus for more critical reflection, and b) to survey preservice teachers to learn if and why they value both types of feedback. The advantages of written feedback from peers are that it is received immediately after observation and it is accompanied by personal and confidential verbal comments from observers. A disadvantage, however, is the veracity of feedback data depends upon the skill and care exhibited by observers. One advantage of a visual record of teaching is that it enables preservice teachers to see and hear their teaching behaviors and conduct a careful critique of their performance. This feedback is not available immediately following the teaching episode however, and its usefulness depends upon preservice teachers' desire to review and reflect upon their teaching. As teacher educators attempt to improve the reflective abilities of preservice teachers, their strategy should include an appropriate stimulus for reflection. Therefore the results of this study are provided to assist educators in informing their choice between peer-generated feedback and visual feedback.

Method

Participants

Participants were preservice teachers enrolled in a junior-level physical education methods course designed to develop effective teaching skills (managerial, instructional, and interpersonal). The course was required for students seeking health and physical education teacher licensure. Preservice teachers systematically

learned skills and then applied them in a variety of laboratory experiences. Two pilot studies and the current reflective study were conducted during three different semesters. In Pilot Study #1, there were nine preservice teachers enrolled (five males, four females), and five were enrolled (three males, two females) in Pilot Study #2. For the current reflective study, 12 preservice teachers were enrolled (six males, six females), including two graduate students seeking teacher licensure.

Materials

Written feedback on participants' teaching performance was provided by two instruments: a rating scale and a coding sheet. The rating scale included a list of teaching skills that were introduced earlier in the course by readings, video and discussion. Observers used the rating scale to report whether participants achieved, partially achieved, or did not achieve each skill. They were also asked to provide a general rating of participants' overall performance and any additional comments. The coding sheet utilized event recording to provide a count of participants' feedback statements, as well as their content, mode, and direction. Earlier in the semester, participants were trained to use both instruments in a live teaching situation. The rating scale and coding sheet were completed either by an observer or the participant.

A mini DVD, purchased by each participant, was used to provide a visual record of teaching. The instructor videotaped participants' teaching using a DVD camcorder and later processed each video. One or two days after teaching, DVDs were returned to participants.

Personal reflections on teaching and teaching feedback were collected on a sheet including two prompts. The purpose of the first prompt ("What's worth doing well?") was to jumpstart the reflective process with a basic reflective question (Hellison & Templin, 1991; van Manen, 1977). The second prompt ("How can you do it better?") encouraged participants to think about their teaching skills and reflect on what they might do to improve them. The questions were purposefully vague to allow stimuli for reflection to emanate from the feedback, not the prompt.

In the current study, a brief survey was used to solicit participant opinions of peer-generated and visual feedback. The survey was administered after all reflections were submitted. It consisted of three open questions: 1) Which type of feedback inspired you to reflect *more deeply* about your teaching? 2) What was more helpful about that type of feedback? 3) Was there any value in the other type of feedback?

Procedures

Pilot study #1. Nine participants were engaged in four peer-teaching labs, each focused on a different teaching skill. They were asked to demonstrate a sport skill, teach and supervise a drill, or teach a game. For the first lab, peer observers gave five participants a completed rating scale or coding sheet and then confidentially discussed their comments. Four participants received only visual feedback and viewed their DVD privately. At the next lab, each group received the other type of feedback. By semester's end, most participants had received equal amounts of both feedback types. All participants utilized the same prompts to reflect upon their teaching performance.

Pilot study #2. In the following semester, five different preservice teachers participated in the same four peer-teaching labs. Procedures for the first two labs were similar to Pilot Study #1, but later, modifications became necessary. Assigning one person to serve as an observer left only four participants to simulate a physical education class. An unusually small class size created an unacceptable teaching situation. Furthermore, it was determined that the rating scale, including a list of teaching behaviors, provided additional reflective prompts. Participants receiving written, peer-generated feedback thus benefited from more reflective prompts than those who received only a visual record of their teaching. Consequently, the

written/verbal feedback component was eliminated for labs three and four. All five participants received visual feedback and used rating scales and coding sheets to reflect upon their own teaching. Due to the small class size, a 75-minute period often provided sufficient time for instructor comments and short discussions after each participant completed a teaching episode. Therefore, unlike labs one and two (and the previous pilot study) when participants received peer-generated feedback or visual feedback only, in labs three and four they all received a visual record, a rating scale or coding sheet, and they experienced an instructor-led discussion of their teaching skills.

Current reflective study. After two pilot studies, procedures were modified for the current reflective study. First, all twelve participants were given rating scales or coding sheets and reflective prompts after each peer-teaching lab. Some participants were observed by a peer, therefore their rating scale or coding sheet contained data and comments from an observer. Others received a blank rating scale or coding sheet and they were free to use them to analyze their own performance while watching their DVD. This provided all participants the opportunity to use the list of teaching behaviors (rating scale) as prompts for reflection. Second, during labs, the instructor made no specific comments about participants' performance. After Pilot Study #2, there was initial concern that instructor feedback given after each participant's teaching episode might have become the subject of their reflection. In essence, they might have been substituting the instructor's reflections for their own. Therefore, in the current reflective study, instructor comments were held until all teaching episodes were complete and then participants were blanketed with thoughts and concepts related to the lab.

Data analysis. In the pilot studies, data analysis was primarily descriptive. Reflective statement length was compared under all feedback conditions and the statements were categorized according to their content. For the current reflective study, data analysis was evaluative in an effort to determine which feedback type inspired reflections that were more critical than technical. A scoring tool was used to evaluate the length, breadth, accuracy and depth of reflections. Length (number of words) was included as a general descriptor. Breadth was the number of topics to which participants referred in their reflection. Accuracy indicated whether or not a teaching concept was referenced correctly. Both breadth and accuracy were weighted more heavily than length because they were indicators of participants' knowledge and understanding of theory introduced in the methods course. The most heavily weighted category was depth because it indicated whether participants exhibited critical thought and penetrated beyond vague conceptual statements to make real connections between learned theory and their performance. This kind of cognitive activity is indicative of critical reflection (Hatton & Smith, 1994).

The data were used to categorize individual reflections into three hierarchical reflective groups: (1) novice, (2) intermediate, and (3) advanced. Furthermore, each participant was given an overall rating of his/her reflective skills. Novice reflections were short (15-33 words) and referenced one teaching concept inaccurately. They exhibited technical language directly from course texts rather than personal thoughts, and made no specific connection between theory and participants' performance. Intermediate reflections were of moderate length (34-52 words), referenced two teaching concepts, and did so with partial accuracy. They included both vague, conceptual statements and specific statements referencing teaching performance. Further, there was a mix of "textbook" language and the participant's own words. Advanced reflections were lengthy (more than 53 words) and addressed three or more teaching concepts accurately. They contained critiques, suggested improvements, and an action plan directly related to participants' performance. Advanced reflections were written

in the participant's own words, exhibiting critical thought and a desire to connect theory with one's performance.

Results

Pilot Studies

The content of reflections submitted in both pilot studies were placed into four descriptive categories: 1) future improvements, 2) critiques, 3) conceptual statements, and 4) "shouldas, or statements reporting what participants should have done. In Pilot Study #1, future improvements were the focus of the majority of reflective statements under both conditions, but those based upon visual feedback were longer (see Table 1). Participants who received visual feedback used a total of 305 words (63%) to describe their future improvements, while peer-generated feedback inspired only 176 words (37%). An example of a future improvement statement was: "One thing I can do better is to slow down and think about what I am going to say." The length of future improvement statements in Pilot Study #2 was nearly identical for both feedback types.

Critique statements reported the existence of a particular teaching behavior and included a self-assessment, such as: "I didn't do a very good job of specific feedback." The length of critiques was similar for both feedback types in Pilot Study #1, but critiques based upon visual feedback were longer in Pilot Study #2. Of the total number of words dedicated to critiques, 93% were inspired by visual feedback. However, it should be noted that in the second pilot study, visual feedback was accompanied by an instructor critique of participants' performance. This may have generated longer critique statements in participants' reflections.

Occasionally, participants would reference teaching concepts in their reflections, such as, "Using routines can save lots of time during managerial transitions." In both pilot studies, peer-generated feedback produced more words assigned to conceptual statements (60% and 71%) than visual feedback. "Shouldas" offered suggestions for what they should have done, such as: "I think I should have paid more attention to their practice and critique them or help them more." Approximately 74% of words allotted to "shouldas" were based upon peer-generated feedback in Pilot Study #1. Since participants received peer-generated feedback on a rating scale listing teaching behaviors related to the lab, they may have reflected specifically upon concepts associated with the list and expressed regret for the behaviors they should have performed.

Current Reflective Study

Reflection levels. In the current study, responses were categorized and scored according to length, breadth, accuracy, and depth. Individual ratings for each reflection were scored and averaged to provide an overall rating of each participant's reflective skills. Reflective skills were distributed normally; two participants were rated as novice, eight were intermediate, and two were at the advanced level. Although a participant was rated at a particular level, his/her individual reflections usually represented more than one level (see Table 2). For instance "LL," an intermediate participant, wrote one advanced and three intermediate reflections. "HWS," wrote two novice and two intermediate reflections, but was categorized as novice. The following are examples of the three levels of reflection.

Novice—"Sharon said I can better my performance by giving more exact feedback while teaching my lesson. Personally I feel I can improve by gaining more confidence in myself and what I am doing and I believe that will come with time." (EL)

This statement was rated as novice because it focused on only one issue (breadth) and it merely restated the peer's opinion, adding vague language about improvement (depth). No connection was made to the participant's performance in the teaching lab. The statement could have been made without reflecting upon peer-generated feedback.

Table 1

Number of reflective words by category and feedback type (Pilot Studies)

Content Category	Peer-generated feedback	Visual feedback
	Pilot Study #1	
Future Improvements	176 (36.6%)	305 (63.4%)
Critiques	70 (49%)	73 (51%)
Conceptual statements	49 (59.8%)	33 (40.2%)
"Shouldas"[a]	43 (74.1%)	15 (25.9%)
	Pilot Study #2[b]	
Future Improvements	65 (50.4%)	64 (49.6%)
Critiques	7 (7.3%)	89 (92.7%)
Conceptual statements	102 (71.3%)	41 (28.7%)
"Shouldas"	0	9 (100%)

[a]Statements reflecting what participants felt they should have done.
[b]Feedback in Pilot Study #2 accompanied by instructor discussion of participants' performance. Visual feedback was accompanied by blank rating scale.

Intermediate—"My managerial task(s) could use work. I felt like I would disperse the group & then remember to tell them things. I need to make sure I go over everything before I disperse." (LL)

Although this statement received low ratings for length and breadth, it accurately referenced managerial theory (accuracy), and it was strongly related to the participant's performance that day (depth).

Advanced—"I need to get closer to the students and teach with confidence & not shy away. I can do the relays in order & at the end do leap frog. That way they are already grouped the entire time & no time would have been lost. I need to start moving around randomly & not be so stationed." (MW)

This statement earned high marks in all four categories. The length was appropriate and the participant addressed three conceptual issues accurately (breadth and accuracy). The reflection focused upon specific events, indicating a mental dissection of the elements of the teaching episode (depth). The statement also included an action plan for future improvement.

Novice, intermediate, and advanced reflections were generated by both types of feedback, but visual feedback inspired more intermediate reflections (61%) from participants at all levels (see Table 2). Among participants at intermediate and advanced levels, nearly an equal number of advanced reflections were written based upon peer-generated and visual feedback (7 and 6 respectively). When advanced and intermediate reflections were combined for all participants, 57% of more critical reflections were inspired by visual feedback. Peer-generated feedback was the source of nearly three-quarters of more technical, novice reflections. Participants at all levels produced more novice reflections based upon peer-generated feedback.

Table 2

Number of reflections by participant rating and feedback type (Current study)

Reflection Rating	Peer-generated feedback	Visual feedback
	All participants (n=12)	
Novice	8 (73%)	3 (27%)
Intermediate	9 (39.1%)	14 (60.9%)
Advanced	7 (53.8%)	6 (46.2%)
	Novice participants (n=2)	
Novice	4 (66.7%)	2 (33.3%)
Intermediate	0	2 (100%)
Advanced	0	0
	Intermediate participants (n=8)	
Novice	3 (75%)	1 (25%)
Intermediate	9 (45%)	11 (55%)
Advanced	4 (57.1%)	3 (42.9%)
	Advanced participants (n=2)	
Novice	1 (100%)	0
Intermediate	0	1 (100%)
Advanced	3 (50%)	3 (50%)

Note. All peer-generated and visual feedback accompanied by rating scale or coding sheet.

Perceptions of feedback. Eleven of 12 participants responded that the visual record of their performance inspired them to reflect more critically about their teaching. Half of them seemed to enjoy examining their teaching and being in control of their critique instead of relying on the opinions of others. Four participants felt that visual feedback was superior to a peer-generated description of their teaching because video was unambiguous and could be viewed repeatedly. This led them to a greater understanding of their strengths and weaknesses as teachers. One interesting and unexpected thought was that viewing a videotaped record of their teaching allowed participants to critique themselves mercilessly. Three participants perceived that their classmates might be "too nice" and that criticism was toned down to avoid hurting their feelings. One participant preferred video because it allowed a critique of her teaching from the perspective of her students.

Despite their preference for visual feedback, the participants also found value in peer-generated feedback. According to five participants, peers provided an extra pair of eyes and ears observing teaching aspects that participants might have

otherwise missed while viewing their videos. Three participants appreciated receiving a perspective different from their own, and two commented that feedback from peers was trustworthy because they were knowledgeable of the teaching assignment as well as their capabilities as a teacher. One participant enjoyed having immediate and personal encouragement from a peer following her teaching episode.

Discussion

A review of results indicates that among preservice teachers in these three studies, a visual record of teaching may have stimulated reflections that were more critical than technical. A total of 14 preservice teachers participated in two pilot studies, and results led to changes in procedures and data analysis for the third and final study. After realizing that the rating scale used for peer-generated feedback provided prompts for reflection in the pilot studies, they were distributed to all 12 participants in the current study, including those who received visual feedback. This provided the structure and guidance recommended for promoting more critical reflections (Crotty & Allyn, 2001; Hatton & Smith, 1994; Langley & Senne, 1997; McCollum, 2002; Tsangaridou & O'Sullivan, 1997; Tsangaridou & Siedentop, 1995.

Rather than describing the length and content of participant reflections, technical reflections in the current study were distinguished from those that exhibited critical thought and connections between theory and practice. Results indicate that after reviewing visual feedback, preservice teachers in the current study wrote 61% of more critical, intermediate reflections. The majority of technical or novice reflections (73%) were produced by peer-generated feedback. Advanced reflections were inspired equally by both feedback types. In this study, preservice teachers who had the desire and ability to reflect critically did so regardless of the type of feedback they received. But for those who required motivation to reflect, visual feedback may have been the impetus to move beyond novice reflections to intermediate reflections. The opportunity to visually review their teaching may have been a stronger stimulus for reflective thought that is more critical than technical.

Preservice teachers in this study overwhelmingly felt that a visual record of their teaching inspired them to reflect more critically. Despite that assertion, they were able to articulate the values of peer-generated feedback. Through periodic observation and feedback from a trusted other, such as a peer, the development of teaching skills was viewed as a cooperative process including the giving and receiving of constructive criticism. There were benefits for both teacher and observer—the teacher receives an alternate perspective of their teaching, and the observer connects theory to practice in a live context. But reviewing a visual record of teaching enabled preservice teachers to see what an observer saw, view it repeatedly, conduct a rigorous self-assessment, and reflect profoundly upon that assessment. Therefore, it is recommended that teacher educators include visual feedback or a combination of both feedback types to inspire critical reflection in their students. Self-assessment of visual feedback may be an effective stimulus to successfully drive preservice teachers to deeper levels of reflection.

REFERENCES

Crotty, T., & Allyn, D. (2001). *Evaluating student reflections*. U.S.: Wisconsin. Retrieved January 9, 2007, from the ERIC database.

Cruickshank, D. (1985). Uses and benefits of reflective teaching. *Phi Delta Kappan, 66,* 704-706.

Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process.* Chicago: D.C. Heath.

Hatton, N., & Smith, D. (1994). *Facilitating reflection: Issues and research.* Australia: New

South Wales. Retrieved January 8, 2007, from the ERIC database.

Hellison, D. R. & Templin, T. J. (1991). *A reflective approach to teaching physical education.* Champaign, IL: Human Kinetics.

Langley, D. J., & Senne, T. (1997). Telling the stories of teaching: Reflective writing for preservice teachers. *Journal of Physical Education, Recreation & Dance, 68(8),* 56-60.

McCollum, S. (2002). The reflective framework for teaching in physical education: A pedagogical tool. *Journal of Physical Education, Recreation & Dance, 73(6),* 39-42.

Napper-Owen, G., & McCallister, S. (2005). What elementary physical education student teachers observe and reflect upon to assist their instruction. *The Physical Educator, 62,* 76-84.

National Association for Sport and Physical Education. (2001). *Guidelines for teacher preparation in physical education: NASPE/ NCATE guidelines* (5th ed.). Reston, VA: Author.

Parkinson, D. D. (2005). Unexpected student reflections from an underused genre. *College Teaching, 53(4),* 147-151.

Risko, V. J., Vukelich, C., Roskos, K., & Carpenter, M. (2002). Preparing teachers for reflective practice: Intentions, contradictions, and possibilities. *Language Arts, 80(2),* 134-144.

Schon, D. (1983). *The reflective practitioner: How professionals think in action.* New York: Basic Books.

Schon, D. (1987). *Educating the reflective practitioner: Toward a new design for teaching and learning in the professions.* San Francisco: Jossey-Bass.

Siedentop, D., & Tannehill, D. (2000). *Developing teaching skills in physical education* (4th ed.). Mountain View, CA: Mayfield.

Sparks-Langer, G. M., & Colton, A.B. (1991). Synthesis of research on teachers' reflective thinking. *Educational Leadership, 48(6),* 37-44.

Tsangaridou, N., & O'Sullivan, M. (1997). The role of reflection in shaping physical education teachers' educational values and practices. *Journal of Teaching in Physical Education, 17,* 2-25.

Tsangaridou, N., & Siedentop, D. (1995). Reflective teaching: A literature review. *Quest, 47,* 212-237.

van Manen, M. (1977). Linking ways of knowing with ways of being practical. *Curriculum Inquiry*, *6,* 205-228.

Zeichner, K. M. (1987). Preparing reflective teachers: An overview of instructional strategies which have been employed in preservice teacher education. *International Journal of Educational Research, 11,* 565-575.

Zeichner, K. M., & Liston, D. P. (1996). *Reflective teaching: An introduction*. Mahwah, NJ: Erlbaum.

Dr. Susan E. King is a faculty member in the Department of Health, Sport, and Exercise Sciences at the University of Kansas, Lawrence, Kansas