DOCUMENT RESUME

ED 385 491 SP 036 109

AUTHOR Van Zoest, Laura R.
TITLE The Impact of Small-Group Discussion on Preservice Teachers' Observations and Reflections.
PUB DATE Apr 95
NOTE 34p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 18-22, 1995).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Audiovisual Aids; *Group Discussion; Higher Education; Mathematics Teachers; *Observation; *Observational Learning; *Preservice Teacher Education; Secondary Education; Teacher Education Programs
IDENTIFIERS Preservice Teachers; *Reflection Process; Reflective Writing; Teacher Thinking; *Video Viewing

ABSTRACT

This study investigated whether providing preservice teachers with opportunities to discuss common observations would have an impact on their observations and reflections. Twelve preservice secondary mathematics teachers in the early stages of their teacher education program were divided into 3 treatment groups based on their orientation towards reflection. All groups observed four video-taped mathematics lessons. Immediately after observing each video, 2 groups spent 50 minutes discussing their perceptions of the lesson. Instead of participating in discussions, the third group observed four additional lessons. Pre- and post-treatment written reflections were collected and analyzed to assess changes in participants' observations and reflections. Results of this analysis support small-group discussion as an effective method for increasing the quantity of reflections. The number of observations did not appear to change but the discussion group participants were better able to recall their observations in post-treatment interviews. (Contains 47 references.) (Author/ND)

THE IMPACT OF SMALL-GROUP DISCUSSION ON PRESERVICE TEACHERS' OBSERVATIONS AND REFLECTIONS

Laura R. Van Zoest

Department of Mathematics
Western Michigan University
Kalamazoo, MI 49008
(616) 387-4527
laura.vanzoest@wmich.edu

Abstract

The study investigated whether providing preservice teachers with opportunities to discuss common observations would have an impact on their observations and reflections. Twelve preservice secondary mathematics teachers in the early stages of their teacher education program were divided into three treatment groups based on their orientation towards reflection. All groups observed four video-taped mathematics lessons. Immediately after observing each video, two groups spent 50 minutes discussing their perceptions of the lesson. Instead of participating in discussions, the third group observed four additional lessons. Pre- and post-treatment written reflections were collected and analyzed to assess changes in participants' observations and reflections. Results of this analysis support small-group discussion as an effective method for increasing the quantity of reflections. The number of observations did not appear to change but the discussion group participants were better able to recall their observations in post-treatment interviews.

Paper presented at the annual meeting of the American Educational Research Association
San Francisco, April 1995

THE IMPACT OF SMALL-GROUP DISCUSSION ON PRESERVICE TEACHERS' OBSERVATIONS AND REFLECTIONS

Laura R. Van Zoest
Western Michigan University

Recent reform documents (National Council of Teachers of Mathematics [NCTM], 1989, 1991; National Research Council, 1989) call for a fundamental restructuring of the way that school mathematics is taught. Achieving this restructuring will require change in the content, method, and assessment of mathematics education (Mathematical Sciences Education Board [MSEB], 1991). Mathematics teachers are seen to be key players in this change:

> The teacher is the gatekeeper to mathematics for our students. What the teacher knows and believes about mathematics, about teaching mathematics, and about the teaching and learning environment determines what students learn and how they will play out their roles as citizens." (MSEB, 1991, p. 17)

The *Professional Standards for Teaching Mathematics (Teaching Standards)* (NCTM, 1991) and *A Call for Change: Recommendations for the Mathematical Preparation of Teachers of Mathematics* (Mathematical Association of America, 1991) focus specifically on the development of teachers able to embody the vision of the reform documents. The authors of these documents recognize the vast range of knowledge and experience that goes into good mathematics teaching and stress the importance of teachers having a strong understanding of mathematics, school mathematics, students as learners of mathematics, and mathematics pedagogy.

Along with university course work, preservice teachers are seen to need a variety of clinical or field-based teaching experiences. These experiences, however, are not to be apprenticeships where the preservice teacher works to imitate the experienced teacher. The *Teaching Standards* states that "teachers should be able to comment and reflect on their own learning environments at the same time that they are involved in clinical and field-based teaching experiences" (NCTM, 1991, p. 125).

The value of reflection to teaching was heralded by Dewey in the early 1900's and echoed by Schon in the 1980's (Borrowman, 1965; Dewey, 1910, 1933; Schon, 1983, 1987). Both of these scholars identified reflection not as a prevalent practice of teachers, but as a means to improve the quality of teaching. In particular, Dewey argued that the field experiences of his time could be characterized as apprenticeships, serving to equip preservice teachers with the technical skills of instruction and classroom management without challenging them to reflect upon their actions (Borrowman, 1965).

Although information about the impact of field experiences on preservice teachers is minimal, the research that has been completed suggests that Dewey's remarks are applicable to teacher education programs throughout history (Borrowman, 1965). Prior to the nineteenth century, field experiences were conceived as apprenticeships and were the only formal education

teachers received (Adler, 1984). As preparation for teaching became institutionalized, field experiences remained central to teacher education programs and retained much of their apprenticeship nature (Conant, 1963).

Dewey advocated classroom observations as a means to provide data for reflection and to lay the foundation for breaking the apprenticeship mode of field experiences (Borrowman, 1965). While contemporary teacher education programs include classroom observations in early field experiences, few have been successful in using them in the manner suggested by Dewey (Adler, 1984). In some cases there are no organized opportunities for reflection. More commonly, students are required to complete an observation form or to write a brief summary for each classroom lesson they observe. These activities, however, do not seem sufficient to induce the depth of reflection advocated by Dewey and the *Teaching Standards*.

A variety of techniques ranging from journal writing to case investigations have been lauded in teacher education for their ability to increase preservice teachers' reflections (Gipe & Richards, 1992; LaBoskey, 1992, 1993; Pultorak, 1993; Richert, 1990; Teitelbaum & Britzman, 1991; Trumbull & Slack, 1991; Zeichner, 1987). Furthermore, entire teacher education programs have been designed around the goal of developing reflective teachers (Korthagen, 1985, 1988; Korthagen & Wubbels, 1991; Valli, 1993; Zeichner, 1990; Zeichner & Tabachnick, 1991). Despite this attention to reflection in the literature, the potential of small-group discussions about common observations to deepen the reflections of preservice teachers seems to have been overlooked.

The isolated nature of early field experiences, even in programs designed to produce reflective teachers, may limit their ability to impact the reflections of preservice teachers. Even in situations where students do discuss what they have observed in the field, the discussions tend to be whole class and more importantly, do not focus on common observations. This deprives preservice teachers of opportunities to discuss and validate their perceptions of the mathematics instruction with others who have observed the same lesson.

The purpose of this study was to investigate whether providing preservice secondary mathematics teachers with opportunities to discuss common observations would have an impact on their observations and reflections. The investigation built on Dewey's notion of observations as sources of data for reflection and was guided by Cognitive Flexibility Theory (Spiro, Coulson, Feltovich, & Anderson, 1988; Spiro, Feltovich, Coulson, & Anderson, 1989; Spiro, Vispoel, Schmitz, Samarapungavan, & Boerger, 1987).

Method

The study compared the observations and reflections of preservice secondary mathematics teachers who participated in small-group discussions after observing classroom mathematics instruction with the observations and reflections of their peers whose early field experience did not include such discussions.

The word "observations" as used in the study refers to what the preservice teachers notice and take mental or written note of when they observe a classroom instructional situation. The word "reflections" as used in the study refers to thoughts about the observations expressed orally or in writing. The study was an experimental investigation that used a pretest-posttest randomized block design (Kirk, 1982).

Instrumentation

Data were collected from five sources: a student information form, a beliefs survey, an orientation-towards-reflection inventory, pre- and post-treatment interviews, and pre- and post-treatment written reflections. The student information form requested background information about the students such as year in school, courses taken, grades, and teaching experiences. The information from this form was used to check for balance in the makeup of the treatment groups.

The pre-treatment interview provided an opportunity for in-depth collection of information about participants' explanations for entering a teacher education program and choosing to major in mathematics. Other questions contained in the interview protocol dealt with participants' past mathematics experiences and what they hoped to learn from various components of their teacher education program.

The main purpose of the post-treatment interview was to debrief the participants at the conclusion of the study. The post-treatment interview protocol (see Figure 1) focused on assessing the participants' recollections and opinions about the lessons they had observed as well as the participants' perceptions of their involvement in the study. Question 8 was included to elicit the participants' views about incorporating small-group discussion about observations into their teacher education program. The post-treatment interview also provided an opportunity to question the participants about ambiguous statements in their written reflections.

Beliefs Survey.

The beliefs survey (Mathematics and Mathematics Teaching Questionnaire) has two components: (a) 15 items addressing beliefs about the nature of mathematics and (b) 25 items concerning beliefs about mathematics teaching and learning. Each of the 40 items was responded to on a five-point Likert scale ranging from Strongly Disagree to Strongly Agree.

The nature of mathematics items were drawn from the Second International Mathematics Study (McKnight et al., 1987). These questions measure perceptions of mathematics as a dynamic process in contrast to perceptions of mathematics as a rigid body of knowledge.

The mathematics teaching and learning items were taken from a survey developed during an earlier study (Van Zoest, Jones, & Thornton, 1994). The questions assess views of mathematics teachers as facilitators of active and social student learning versus views of mathematics teachers as directive providers of information to passive students. For 186 preservice teachers invoived in a pilot for the previous study, the Cronbach's alpha for the teaching and learning component of the beliefs survey was .81.

The beliefs survey was used to collect background information on the participants' views towards mathematics and towards the teaching and learning of mathematics. Along with identifying potential biases among the groups, it was possible that knowing the students' perspectives could provide explanatory information about their observations and reflections.

According to the results of the beliefs survey, the participants could be described as weakly agreeing with the view of mathematics teachers as facilitators of active and social mathematics learning ($N = 12$, $M = 3.70$, $SD = 0.19$). Preservice teachers in an earlier study (Van Zoest, Jones, & Thornton, 1994) expressed stronger agreement ($N = 103$, $M = 4.09$, $SD = 0.29$), possibly because they were further along in their teacher education program. Unlike the students in the earlier study, the students in the current study had not been directly exposed to the ideas of the National Council of Teachers of Mathematics (1989) *Curriculum and Evaluation Standards for School Mathematics* prior to completing the beliefs survey.

Internal/External Orientation Test.

The internal/external orientation test (I.E.O.-Test) was developed by Korthagen and his colleagues while assessing the Stichting Opleiding Leraren (SOL) secondary mathematics teacher education program in the Netherlands (Korthagen, 1988; Korthagen & Wubbels, 1991). Their initial research led them to believe that some students were disposed in favor of training in reflective teaching while others were not. A similar conclusion was reached by LaBoskey (1989) in her study of the stability of reflection in preservice teachers.

The I.E.O.-Test for students in the initial stages of their teacher education program was designed to assess orientation towards reflection in three domains: (a) the preservice teacher him or herself, (b) his or her fellow students, and (c) mathematics (Korthagen, 1988). For each domain, a scale for internal learning orientation and a scale for external learning orientation was developed. Students who rate high on the three internal scales tend to depend on themselves for direction and would be more likely to thrive in situations that encourage reflective thinking. Students who rate high on the three external scales tend to be more comfortable in a directive environment.

When the Dutch version of the I.E.O.-Test was administered to 138 first- and second-year students from three teachers' colleges in the Netherlands, the Cronbach's alphas for the six scales ranged from .77 to .87 (Korthagen, 1988). The I.E.O.-Test was translated to English for the current study. During the translation, the researcher was in communication with the developers of the original instrument to verify the accuracy of each item. The English version was then administered to a group of 19 preservice secondary mathematics teachers at the University where the study took place. Despite the small sample size, the Cronbach's alphas ranged from .74 to .95, in line with the reliabilities calculated for the Dutch version.

The I.E.O.-Test was used in the current study to control for orientation towards reflection as a factor accounting for changes in reflection during the treatment. The randomized blocks were formed based on scores from this inventory. Consultation with the developers of the I.E.O.-Test

resulted in the decision to sum the participants' scores on the internal scales to arrive at one score measuring orientation towards reflection. This appeared to be the most efficient and effective way to rank order the participants regarding their natural inclination to reflect. Once the 12 participants were rank ordered, they were divided into four levels of reflection: low, low-medium, medium-high, high. Each level contained three participants. Within each level the participants were then randomly assigned to the three treatment groups: (a) observation-only, (b) observation-with-discussion, (c) observation-with-facilitated-discussion.

Participants

The participants in the study were drawn from a pool of 16 secondary mathematics education majors enrolled in two mandatory professional education courses. All of the students were invited to participate and were offered clinical hours for time spent in the study. Within one week of the invitation, 13 of the students agreed to participate. All were accepted, although the optimal number of participants for the study was 12. One of the participants dropped the course the following week and hence was excluded from the study. Of the three non-participants, one immediately declined to participate, one volunteered to participate at a point too far in the data collection to be included, and the other dropped out of the course. The final sample included 12 preservice secondary school mathematics teachers at the beginning of their professional education course work.

As is typical at the University, the participants were white and middle-class. Table 1 provides a focal point for the following discussion of selected participant characteristics. Three of the students were male. Two of the participants were non-traditional students in the sense that they were returning for their teaching credentials after pursuing other interests. All but one of the participants were transfer students. Although that is a high percentage, it was not unusual. The University draws students who attend community colleges for the first year or two of their studies, students who leave private schools for financial reasons, and students who transfer from other state universities when they decide to enter teaching. All three types of students were represented in the study. Even though half of the participants had achieved senior status, all had two full semesters beyond the semester of the study remaining in their teacher education program.

The individual participants had taken between 6 and 10 mathematics content courses and their grade point averages for these courses ranged from 2.20 to 3.86. Three had worked as undergraduate teaching assistants (UTAs) in the Department of Mathematics at the University. As UTAs, they led discussion sessions, held office hours, and kept records for a college algebra course. The three UTAs and one additional participant had tutored college students in the Department's mathematics assistance laboratory.

Random assignment of the participants to the treatment groups was based on the sum of their scores on the internal scales of the I.E.O.-Test. Kruskal-Wallis one-way analysis of variance by ranks (Siegel & Castellan, 1988) conducted on the internal score as well as mathematics grade

point average, number of mathematics courses taken, and scores on the two components of the beliefs survey yielded no significant differences among the three treatment groups.

The mean internal score on the I.E.O.-Test for all 12 participants was 3.5. That is the same as the mean of 10 Dutch preservice secondary mathematics teachers involved in a longitudinal study (Korthagen & Wubbels, 1991). The range for the students in the current study was slightly wider (3.1-4.4) than that of the Dutch students (3.2-4.1). The means for the observation-only, observation-with-discussion, and observation-with-facilitated-discussion groups were 3.5, 3.6, and 3.4, respectively.

Although there was no statistical test available for categorical variables in a sample of this size, inspection of the groups reveals reasonably well-balanced groups. Each group contained one male, a mix of juniors and seniors, and at least one person with some form of teaching experience. There were, however, two differences that may affect conclusions drawn from the study. Both of the non-traditional students were assigned to the observation-only group and all the observation-with-facilitated-discussion group members were enrolled in the second of the two professional education courses.

Procedures

The first phase of the study involved the collection of background data and formation of the treatment groups. Upon agreeing to participate in the study the students scheduled a 30-minute meeting with the researcher. The meeting began with the students taking approximately 10 minutes to complete the beliefs survey and orientation-towards-reflection inventory. The researcher then interviewed the students using the pre-treatment interview protocol. The interviews were audiotaped and subsequently transcribed. The responses to the I.E.O.-Test were analyzed and, based on the analysis, participants were assigned to four levels of reflection. Participants from each level were randomly assigned to the three treatment groups: (a) observation-only, (b) observation-with-discussion, and (c) observation-with-facilitated-discussion.

All participants then observed the pre-treatment videotaped mathematics lesson. Immediately after viewing the video they were given the following instructions:

> Please spend the next 30 minutes writing about what you observed in this lesson. Include as much information as possible and be sure to add your own views of what occurred during the lesson. You may use any notes that you have but please do not talk to each other.

The participants had also been shown these directions prior to watching the videotaped lesson and were provided with paper and pencil for taking notes. The written reflections were collected, typed, and entered into NUD•IST (Replee Proprietary Limited, 1993). The participants' notes were also collected, but upon examination it was discovered that they did not add any information to the written reflections.

After the written reflections had been completed the treatment phase of the study began. Over the course of the next three weeks the three treatment groups observed the sequence of four

videotaped high school mathematics lessons described in the next section. The observation-only group observed an additional four live mathematics lessons of their choosing.

Immediately after watching each videotape, the discussion groups spent 50 minutes discussing their perceptions of the lesson. The discussions were videotaped for later analysis. Each member of the observation-with-discussion group led the group's discussion during one of the four sessions. The discussion leader for the session was provided with a stack of discussion questions (see Figure 2) and given the following directions:

> Make sure to address all the questions in the 50 minutes. If you finish early, go back and discuss some questions in greater depth or address other aspects of the lesson that interested you.

The researcher brought the discussion cards to the observation-with-discussion group at the beginning of each discussion session and then returned to collect them at the end of the allotted 50 minutes.

The observation-with-facilitated discussion group differed in that the researcher instead of a participant led the discussion sessions. The same questions were addressed, but the order in which they were asked was at the discretion of the researcher. In general, this did not differ from the order of the stack of questions provided to the observation-with-discussion group. The researcher did not participate in the discussion other than to ask open-ended, non-evaluative, probing questions. The participants' inquiry "When are you going to tell us what you think?" provided some evidence that the researcher's attempt to remain a neutral force was successful.

After the three groups had completed the treatment phase of the study, the participants observed the post-treatment video and completed a written reflection under the same circumstances as with the pre-treatment video. The final participant involvement with the study was the post-treatment interview. Answering the questions on the post-treatment interview protocol took most participants about 20 minutes. Responses to the last question "Is there anything else you would like to add?" and follow-up probes caused many interviews to last another 10 to 15 minutes. The longest interview was 90 minutes. The interviews were audiotaped and subsequently transcribed.

Videotaped Mathematics Lessons

For this study, the classroom observations were videotaped classroom mathematics lessons. Selection of the videotaped lessons was guided by the tenets of Cognitive Flexibility Theory (Spiro et al., 1988; Spiro et al., 1987). That is, the videos were chosen in such a way that each lesson had both elements in common with adjacent lessons and elements that differed.

Extremely uneventful or poorly managed classes were excluded from consideration. Observing classes that did not have much activity either by the teacher or the students would have provided little stimulus for reflection. In a similar manner, having the students observe poorly managed classes was seen to interfere with the purposes of the study. In an earlier study the researcher found that observing serious classroom control problems overwhelmed preservice teachers and caused them to focus on management issues to the exclusion of other equally relevant

aspects of the lesson (Van Zoest, 1993). In the interest of having the students benefit mathematically as well as pedagogically, the mathematics content in the selected videos satisfied at least one of the following criteria: (a) typically taught by first-year teachers, (b) relatively new to the high school curriculum (NCTM, 1989), or (c) commonly misunderstood by preservice teachers (e.g., Ball, 1990; Ebert, 1993; Leinhardt, 1989).

The pre- and post-treatment videos were chosen to be representative of "typical" high school mathematics classroom lessons. To minimize extraneous influences these videos were of consecutive lessons taught by the same teacher to the same class. The lessons were primarily teacher-directed with student involvement limited to asking and answering questions. The teaching strategies were reasonably varied and reference was made of relevant history and real-world applications. Lecture was interspersed with seat work during which the students were allowed to work together. The pre- and post-treatment videos were most likely to resemble the mathematics instruction that the preservice teachers had experienced in their own education (Ball, 1990).

The treatment videos, on the other hand, were chosen, in line with the tenets of Cognitive Flexibility Theory, to "criss-cross" the domain of secondary school mathematics as much as was possible with four lessons. A larger number of observations would have covered more territory, but based on the feedback from an earlier study it was determined that four was the optimal number to provide the maximum variety without overtaxing the participants' schedules and endurance (Van Zoest, 1993). In the earlier study the benefits of observations followed by discussion were considered to peak at about three sessions and then begin to diminish.

Data Analysis

The pre- and post-treatment written reflections formed the core of the data to be analyzed. Other data sources, such as the Student Information Form, Mathematics and Mathematics Teaching Questionnaire, I.E.O.-Test, and the pre- and post-treatment interviews, were used to provide a context within which to interpret the participants' responses to the treatment. Additional sources allowed the researcher to guard against incorrect conclusions through triangulation of the data (Eisenhart, 1988). As the additional data sources have already been explained, the following discussion focuses on the procedure used to analyze the written reflections.

The first step was to define the unit of analysis. It was decided that the most appropriate unit of analysis for the study would be a coherent statement or partial statement that described either an observation or reflection. These were referred to as "thoughts." In a few cases one thought contained both an observation and a reflection. If separating the two would sacrifice coherency, the thought was coded as both an observation and a reflection.

After each participants' pre- and post-treatment reflections were broken down into a sequence of thoughts, the documents were entered into NUD•IST. The thoughts were then assigned multiple-level codes. The first tier identified whether the thoughts were from a pre- or post-treatment written reflection. The second tier labeled them as either an observation or a

reflection. As mentioned earlier, there were some instances when a thought was coded as both an observation and a reflection.

If a thought was coded as a reflection, it was also given a type code. A reflection rubric developed through a grounded theory approach (Strauss & Corbin, 1990) in the context of the earlier study (Van Zoest, 1993) formed the basis for the type codes. During the process of refining the codes a series of written reflections were "double-coded" (Miles & Huberman, 1984, p. 60). That is, the researcher and other mathematics educators first coded the written reflections independently and then discussed differences until consensus was reached. The modified types can be classified into two categories: (a) reflections without justification, and (b) reflections with justification. Three of the modified types fell under the category of reflections without justification: (a) general reflections, (b) specific reflections, and (c) personal reflections. The other two modified types fell under the category of reflections with justification: (a) reflection justified by experience and (b) reflection justified by logical reasoning. The personal reflections were comments made by the preservice teacher about him or herself. Examples of personal reflections include "A big concern of mine is to determine how to motivate students in my classroom" and "This is a very interesting example that I will be sure to use in my class someday."

Figure 3 contains examples of the types, excluding personal reflections. In a loose sense the types of reflection were seen along a continuum of robustness. Three examples, one positive, one negative, and one neutral, are included in Figure 3 to provide a sense of the building of robustness across the continuum. A specific reflection justified by logical reasoning would likely be more robust than a general reflection without justification. The highest level of robustness would be reserved for reflections based on logical reasoning that drew on professional knowledge. Although possible, students at the beginning of their teacher education program would not be expected to have the knowledge base necessary to engage in reflection at the highest level. Therefore it was not surprising that the highest level of robustness was not required in coding the preservice teachers' written reflections.

No traditional measure of inter-rater reliability seemed appropriate to the current study. Each written reflection was coded a minimum of three times with any differences resolved through reviewing the definitions of the codes. Additionally, thoughts that received the same code were compared to insure that they were similar enough to be grouped together under one label. This activity both provided an error check and assisted in refining the code definitions. In general, the use of an iterative approach to coding tightened the code definitions and provided multiple checks for accuracy.

At the completion of the coding the number of thoughts in each category was counted and analyzed using SPSS (SPSS, Incorporated, 1990). The categories were searched for patterns and movement of the groups relevant to their initial positions. Line plots were used to provide graphical representations of the data. The Kruskal-Wallis one-way analysis of variance by ranks (Siegel & Castellan, 1988) was used to test for significant differences among the treatment groups'

number of thoughts, observations, and reflections. For the cases that had significant differences, follow-up multiple comparison tests as provided by Siegel and Castellan (1988) were performed to determine which groups differed.

Results

Number of Thoughts, Observations and Reflections

Means and standard deviations for the pre- and post-treatment number of thoughts, observations, and reflections were calculated for each of the three treatment groups: (a) observation-only, (b) observation-with-discussion, (c) and observation-with-facilitated-discussion. For number of thoughts, observations, and reflections, the means for the observation-only group decreased but the means for the observation-with-discussion and observation-with-facilitated-discussion groups increased or stayed the same (see Table 2).

To determine whether these differences were statistically significant, the Kruskal-Wallis one-way analysis of variance by ranks (Siegel & Castellan, 1988) was first applied to the number of thoughts, observations, and reflections in the participants' pre-treatment written reflections. This was done to determine whether the pre-treatment scores needed to be taken into account in the analysis of the post-treatment scores. Due to the low power present with a small sample size, the level of significance was set at .10. None of the differences were significant at this level. This fact suggested that the random assignment had been successful and that it was not necessary to control for the participants' pre-treatment scores on these variables. The Kruskal-Wallis test was then applied to the number of thoughts, observations, and reflections from the post-treatment written reflections. Table 3 contains the groups' mean ranks on these variables.

Table 4 contains the chi-square approximations of the Kruskal-Wallis distributions for the number of thoughts, observations, and reflections in the participants' post-treatment written reflections (Ferguson, 1971). There were significant differences among the groups for the post-treatment number of thoughts and reflections but not for the number of observations. For the cases that had significant differences, follow-up multiple comparison tests as provided by Siegel and Castellan (1988) were performed to determine which groups differed. The significant comparison for the number of thoughts was between the observation-only and the observation-with-discussion groups. The observation-with-discussion group had significantly more thoughts on the pre-treatment reflection than did the observation-only group. Both discussion groups had significantly more post-treatment reflections than the observation-only group but did not differ significantly from each other.

Based on this analysis, it appears that small-group discussion increased the number of reflections for both facilitated and non-facilitated discussion groups but increased the number of thoughts only for the observation-with-discussion group. It is important to note that due to the small sample size the statistical tests had low power. That is, one can be fairly confident that statistically significant differences are true differences but must be skeptical of accepting equality when the differences are not statistically significant. Siegel and Castellan caution that in these

cases the researcher should "seek corroborating evidence or obtain additional data" (1988, p. 210). The auxiliary sources of data for this study such as the pre- and post-treatment interviews served that purpose and are integrated into the discussion section of the paper.

Reflection Types

Also of interest was whether involvement in small-group discussions about classroom observations would impact the robustness of participants' reflections. To address this question, the pre- and post-treatment written reflections for the three groups were examined for movement in reflection types. Table 5 contains the groups' means and standard deviations for five types of reflection: (a) general without justification, (b) specific without justification, (c) personal, (d) justified by experience, and (e) justified by logical reasoning.

After the reflections were divided into types, the numbers of each type were too small to benefit from statistical tests for significant differences among the groups. Line graphs provided a visual aid for identifying movement among the mean pre- and post-treatment reflections for the following types: general reflections without justification (Figure 4), specific reflections without justification (Figure 5), and reflections justified by logical reasoning (Figure 6). Because the personal and justified-by-experience types contained so few reflections, they were excluded from further analysis.

Both discussion groups had an increase in their mean numbers of reflections categorized as general reflections without justification and specific reflections without justification. The observation-only group, on the other hand, experienced a decrease in their mean number of reflections for both of these types. The results in the category of reflections justified by logical reasoning were mixed. For this category, the observation-with-facilitated-discussion group decreased, the observation-with-discussion increased, and the observation-only remained the same. Both the increase and the decrease of the respective discussion groups were less extreme than all but one of the changes in the reflection-without-justification categories.

Earlier an increase in the number of reflections for both discussion groups was identified in relation to a corresponding decrease for the observation-only group. Based on the analysis of movement among reflection types, it seems that the discussion groups' increase was in reflections categorized as general or specific without justification. In other words, it does not appear that the small-group discussion affected the robustness of the participants' reflections. Instead, it seems that small-group discussion increased the number of reflections at the lowest levels of cognitive involvement without moving reflections to higher levels of robustness.

Discussion

Observations and Reflections

The perception of the researcher, based on observation of the discussion sessions and the post-treatment interviews, was that involvement in the discussion groups was more beneficial to the participants than was evident in the written reflections. One possible explanation for the discrepancy is that the pre- and post-treatment videos were not rich enough to allow an increase in

observations or reflections between the pre- and post-treatment writing sessions. As suggested by Dewey (1910, 1933), the classroom observations were to provide data for reflection. The students may have noticed and reflected on the bulk of the data present in the pre-treatment video in their initial written reflection. Because the post-treatment video was chosen to be similar to the pre-treatment video, the students would then only be able to repeat themselves in the post-treatment written reflection. The consistent focus of the observations supports this scenario.

Another explanation is the limitation of written reflections. Writing about an observation may only reproduce a segment of the data and preliminary reflections. This is consistent with Pultorak's (1993) finding that higher levels of reflective judgment were more apparent in oral interviews than in written reflections. Pultorak hypothesized that the difference may be due to the interviewer's ability to probe for deeper insights. In a similar manner, the social interaction and influx of ideas that occur in a small-group discussion with peers may foster more robust reflection and lasting recollections of the observed lesson than can be captured in a written reflection.

The researcher's view that students in the discussion groups had a richer experience was based in part on the students' abilities to discuss the various videos (see Table 6). The students in both discussion groups illustrated their statements during the post-treatment interview with details from the videos. In the observation-only group, two of the students seemed to have a difficult time even recalling general content of the videos. Furthermore, three of the observation-with-discussion group students, Bob, Erin, and Nan, had vivid recollections of specific incidents and became animated when speaking about the incidents during their post-treatment interviews. Fran, the fourth member of the observation-with-discussion group, explained during her post-treatment interview that her poor eyesight had made catching the details of the lesson from the T.V. screen difficult. She relied heavily on the discussion to pick up the details of what had occurred in the classrooms. The dynamics of the discussion groups and the preservice teachers' perceptions of involvement in small-group discussions form another dimension of the study.

The Discussion Groups

The observation-with-facilitated-discussion group generally found it difficult to discuss each video for a full 50 minutes. The planned questions had usually been exhaustively addressed within 30 minutes. In the interest of maintaining consistent discussion times, the facilitator used the remaining time to ask additional questions that were tangentially related to the lesson. One member of the group, Rand, who was undecided about the value of discussion over additional observations, suggested that the ideal would be to increase the number of videos and reduce the discussions to 30 minutes. It is possible that a reduced discussion time would have been fully as effective for this particular group of students.

The observation-with-discussion group provided a direct contrast. Each time the researcher alerted them that the 50 minute discussion period was over they were still discussing, sometimes heatedly. It was not uncommon to walk past the group's meeting room and hear group members arguing or excitedly chiming in agreement. That kind of spontaneous conversation rarely occurred

in the facilitated group. Whether the cause was the presence of the facilitator or the personalities of the students is less clear. Nan, a member of the observation-with-discussion group, seemed to think that the lack of a facilitator was a key to the success of their group:

> Having someone else ask the questions, I think may have gotten in the way of pure discussion like we had, because the only people that were there was just us and the video camera. And we pretty much managed to ignore the video camera and we didn't feel we had to watch out for anybody but each other. If someone else had been reading the questions, it would have gotten in the way of the discussion. (Nan, post-interview)

Erin was more specific: "I think people might not have said everything because they would have thought that this is not what she wants us to say so I better not say it" (Erin, post-interview). One area in which there was evidence of this was in the language used in the discussions. The observation-with-discussion group used more explicit descriptive language, including profanity, than the observation-with-facilitated-discussion group. This frankness suggests that the non-facilitated group was more comfortable and relaxed during their discussions. The presence of a facilitator may have formalized the social interaction.

Another difference between the two discussion groups appeared in the participants' responses to the first question of the post-treatment interview: "What about the experiences you've had in this study stands out most in your mind?" The participants' initial responses are provided in Table 7. It is striking that each person in the observation-with-discussion group mentioned something related to their involvement in the small-group discussion. The observation-with-facilitated-discussion group, on the other hand, first referred to aspects of the observations. In this regard, the observation-with-facilitated-discussion group was more similar to the observation-only group than to the observation-with-discussion group. One reason for this may have been the stronger group dynamic in the observation-with-discussion group.

Within the observation-with-discussion group, Bob and Erin were very confident in their opinions and, by nature, seemed to want to share them with others. Fran, on the other hand, spoke infrequently. When she did share her opinions, she was often in disagreement with Bob and Erin. In these situations Nan tended to act as a mediator, relating to both sides of the issue and assisting Fran in expressing her view. Fran explained why Nan needed to intercede to maintain the momentum of the discussion: "I don't know how to express my own ideas enough to tell them why I think they're wrong" (Fran, post-interview).

In contrast, all the members of observation-with-facilitated-discussion group appeared content to form their own opinions without convincing others that they were right. The most controversial point in the discussions occurred in response to the pre-calculus teacher's (Treatment Video 4) use of graphing calculators. Gina, Lynn, and Mary expressed concern about using technology in their classrooms, and were unclear about the role graphing calculators and computers should play in mathematics instruction. Rand, on the other hand, was very comfortable with both technologies and felt they were important teaching tools. During the discussion he stated his view

but made little effort to challenge the other group members' perceptions. The facilitator asked questions designed to encourage further exploration of the issue, but the group seemed reluctant to examine the topic in any depth. In general, the observation-with-facilitated-discussion group seemed content to state their opinions and to stop the conversation at that point.

Lynn's comment may summarize another problem with the observation-with-facilitated-discussion group: "We all pretty much liked and disliked the same things. All of us pretty much had a consensus about what we thought was good teaching. I think we all want to teach almost the same way" (Lynn, post-interview). The fact that all four members of the observation-with-facilitated-discussion group were enrolled in the second of the two professional education courses (see Table 1) may have limited the variety of their contributions. Gina's comment, "I would say that most of our ideas are the same because we're going through the same experiences right now" (Gina, post-interview), supports that possibility. The effectiveness of case discussions with inservice teachers (Barnett, 1991; Barnett & Sather, 1992) may be due in part to the variety of experiences that the teachers brought to the discussions. The majority of preservice teachers may lack the background necessary for rich conversation about instructional situations.

According to responses during the post-treatment interviews (Question 8), six of the eight participants in the discussion groups would participate in observations followed by discussions again if they were given the opportunity (see Table 6). Rationale given by these participants for choosing the discussion option included the following: more interesting, get ideas from others, reinforces what observed, pay more attention to lesson, notice things that might have otherwise missed, and find out what other people think is important. Tom, a member of the observation-only group, gave similar reasons for his decision:

> I think you can miss things and only look at certain situations. If you bring other people's views in that strengthens your own opinion. That's why I think it would be helpful to talk to a group about the lessons. Plus it might make you pay more attention. Because you know you're going to talk about it. (Tom, post-interview)

In contrast, another participant in the observation-only group, Dora, preferred more observations to observations followed by discussion. Her reason reflects a focus on acquiring practical skills for teaching:

> In observing you can at least get ideas from teachers who are actually out there doing it and succeeding at least enough to keep their job, whereas students, we can come up with all kinds of ideas but we don't know if they'll work. It's the blind leading the blind, I think. (Dora, post-interview)

Although she also carefully considered what she observed, in some sense Dora had been in an apprentice mode prior to beginning her teacher education program: "Since I was in high school I knew I wanted to be a teacher so I really started trying to watch my teachers and see how they handled things and I had some pretty good examples" (Dora, pre-interview). Dora believed strongly in the value of discussing observations but she felt that any organized discussion should

be with the teacher of the observed lesson or some other inservice teacher who could bring extensive experience and professional knowledge to the discussion.

Val, the member of the observation-only group who was undecided about which option she would choose, mentioned her plans to discuss the videos with Dora after the study was over. She had found the researcher's request to not discuss her experiences until the end of the study a difficult one to keep and was looking forward to being able to talk to the other participants. Some students will naturally talk about what they have observed with friends who are often also preservice teachers.

Students such as Val and Dora who already participate in informal small-group discussion are not the students who would benefit the most from the integration of small-group discussion into the observation component of early field experiences. Instead, the key beneficiaries would be the students who for some reason are isolated from their preservice-teacher peers or who lack the motivation to become acute observers of classroom instruction without outside assistance.

The small-group discussions were most effective for reinforcing and supplementing what the participants could have learned from watching the video alone. Although the discussion did not seem to radically alter the participants' perspectives, it did seem to provide the preservice teachers with the tools to better articulate their initially held views. This is consistent with Zeichner and Liston's (1987) findings regarding student teaching.

The success of a student in small-group work is another factor to be considered. Trumbull and Slack (1991) reported on a possible connection between reflective development and successful small-group work. Although none of the participants would be considered unsuccessful in the small-group discussion, two participants, Fran and Rand, were less active than their peers. Additionally these two were the only participants in the discussion groups to not express a clear preference for small-group discussion over additional observations (see Table 6). Fran preferred totally eliminating the discussion component of the observations for the following reason:

> It's just other people's opinions and most of the opinions that were expressed I disagreed with, but I'm not a very good arguer so I don't tell people I disagree with them. I like knowledge, I like facts and what other people say always confuses me. (Fran, post-interview)

Rand did not want to eliminate the discussion component completely because he felt that there was some benefit: "I guess it helped to clarify things in my mind and it was interesting to see what other people thought" (Rand, post-interview). He also thought that devoting 50 minutes per lesson to discussion as in the current study was too much and suggested a compromise of more videos and 30-minute discussions.

If there was a connection between reflection development and small-group work, Rand and Fran would be expected to have a low number of pre- and post-reflections and to have low internal scores on the I.E.O.-Test. This was true for Rand. He had the lowest pre-treatment and post-treatment reflection counts of participants in the discussion groups and the lowest internal score of

the sample on the I.E.O.-Test. Fran, on the other hand, had the highest post-reflection count and the fourth highest overall internal I.E.O.-Test score.

Although Fran seems to provide a counterexample to the existence of a connection between reflection development and small-group work, closer inspection of her data reveals that this is not the case. First of all it must be remembered that the I.E.O.-Test has three components: self, fellow students, and mathematics. Fran had the lowest score of the sample on the fellow-students component of the scale, the component most likely to reflect orientation towards successful group work. Yet this was masked in Fran's case by high scores in the areas of self and mathematics.

Additionally, six of Fran's post-treatment reflections were personal. That is, reflections about her own thoughts or concerns and not the classroom instruction. If the personal reflections were removed from the count, the number of her post-treatment reflections would fall at the total group's median. The data for Fran do not provide a counter example but instead lend moderate support to the idea of a connection between reflection and successful group work. In general, then, the study results provide some corroboration for a connection between less-successful small-group work and reduced levels of reflection.

Limitations

A key limitation of the study was that the written reflections may not have been able to measure the increase in reflection caused by involvement in small-group discussion sessions. The process of writing itself may affect reflection (Gipe & Richards, 1992; Pultorak, 1993) and may not capture the nature of reflection generated by small-group discussion.

Furthermore, at least some of the participants observed differently because they were going to be writing: "When I had to write for the first and last one, I paid more attention, or at least I thought I paid more attention. I was thinking about having to write something" (Sara, post-interview). The instructions for the pre- and post-treatment written reflections informed the students that they were going to be asked to write at the end and encouraged them to take notes. There was no parallel encouragement provided to any of the groups during the treatment phase of the study.

Another limitation was the small and homogeneous sample. Multiple groups for the three treatments would have provided insight into the impact of group dynamics and individual personalities on the small-group discussions. Because the students in the sample were white and middle-class, the results are not automatically generalizable to the larger population of preservice teachers. This is particularly unfortunate in light of recent attempts to recruit minorities into the teaching field. Furthermore, it is possible that small-group discussions would be more beneficial to minorities and that the discussions would be richer due to a larger range of perspectives.

The students in the videos were also primarily white and middle-class, limiting the opportunities for equity and justice issues to arise during the discussions. Neither the videos nor the composition of the groups provided a stimulus for discussing the multicultural issues that U.S.

teachers will increasingly face as whites move toward being a minority of the nation's population (National Research Council, 1989).

The use of both quantitative and qualitative data in the study was not as rich as it might have been. The sample size of 12 was small for a quantitative study yet large for gathering the extensive in-depth data required of a qualitative study. The compromise resulted in decreased power for the quantitative component and decreased description for the qualitative component.

Conclusions & Implications

The current study was designed to assess the value of small-group discussions for improving preservice secondary mathematics teachers' early field experiences within the confines of the current system and without incurring additional costs. The results of the study indicate that small-group discussions may hold a key to unlocking the benefits of classroom observations, particularly for students who are isolated from their preservice-teacher peers or are not sufficiently motivated to become acute observers of classroom instruction without outside assistance. Small-group discussions about shared observations provide preservice teachers with a common ground on which to base their communication. Integrating non-facilitated small-group discussions into the observation component of early field experiences is recommended for the following reasons:

1. Increasing reflections at the low end of the robustness scale lays a foundation on which later experiences can build.
2. The students seem to pay more attention to the lessons if they know they are going to discuss them later, therefore gaining more from the observations.
3. Discussion groups provide an opportunity for preservice teachers to develop collegial relationships with their preservice colleagues.

Videotaping of the discussions is encouraged for two reasons: (a) to hold the students accountable for completing the requirement, and (b) to provide the course instructor with a means to gain insight into the perceptions of his or her students.

Although further research is needed to verify the benefits of small-group discussion, there is no evidence that the discussion would be detrimental to the preservice teachers. Any initial scheduling difficulties would be overshadowed by the potential benefits. No added personnel would be necessary and no additional time would be required of the students. Furthermore, if the university already has microteaching facilities equipped with videotaping equipment no capital investment or rental fees would be required for videotaping the discussions.

Recommendations for Further Research

This study was a first step in the exploration of the relationship between small-group discussion about common observations and the participants' observations and reflections. Further research is needed both to verify and extend the findings of the current study. Recommendations for further research based on this study include the following:

1. The current study used entire class periods for the observations. It is possible that shorter segments containing critical incidents or cases of specific aspects of classroom instruction

would provide richer sources for discussion. Any classroom lesson contains "down time" and repetitive actions on the part of both the teacher and the students. Reducing a classroom lesson to a shorter segment while still capturing its salient characteristics may prove to be a more efficient use of students' observation time.

2. The personal dynamics within the discussion groups of this study seemed to affect the nature of the group's discussion. Replicating the current study would provide some information about the impact of personalities on the success of small-group discussion. Forming the groups homogeneously by level of reflection would provide additional information.
3. The students both in the study and in the video were primarily white and middle-class. Investigation into the effects of small-group discussion involving more diverse populations is needed.
4. Alternative methods of assessing the students' observations and reflections need to be investigated. Possibilities include group interviews, individual interviews, and personal tape recordings. For the latter, students would "speak" their observations and reflections rather than write them. The informal nature of speech, when compared to writing, may more closely capture the type of reflection generated by small-group discussion.

REFERENCES

Adler, S.A. (1984, April). *An historical analysis of early field experiences*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Ball, D.L. (1990). The mathematical understandings that prospective teachers bring to teacher education. *The Elementary School Journal, 90*, 449-466.

Barnett, C. (1991). Building a case-based curriculum to enhance the pedagogical content knowledge of mathematics teachers. *Journal of Teacher Education, 42*(4), 1-9.

Barnett, C., & Sather, S. (1992, April). *Using case discussions to promote changes in beliefs among mathematics teachers*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Borrowman, M.L. (Ed.). (1965). *Teacher education in America: A documentary history*. New York: Teachers College Press.

Conant, J.B. (1963). *The education of American teachers*. New York: McGraw-Hill.

Dewey, J. (1910). *How we think*. New York: D.C. Heath.

Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process* (2nd ed.). Boston: D.C. Heath.

Ebert, C. L. (1993, April). *An assessment of prospective secondary teachers' pedagogical content knowledge about functions and graphs*. Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA.

Eisenhart, M.A. (1988). The ethnographic research tradition and mathematics education research. *Journal for Research in Mathematics Education, 19*, 99-114.

Ferguson, G.A. (1971). *Statistical analysis in psychology and education* (3rd ed.). New York: McGraw-Hill.

Gipe, J.P., & Richards, J.C. (1992). Reflective thinking and growth in novices' teaching abilities. *Journal of Educational Research, 86*, 52-57.

Kirk, R.E. (1982). *Experimental design* (2nd ed.). Belmont, CA: Brooks/Cole Publishing Company.

Korthagen, F.A.J. (1985). Reflective teaching and preservice education in the Netherlands. *Journal of Teacher Education, 36*(5), 11-15.

Korthagen, F.A.J. (1988). The influence of learning orientations on the development of reflective teaching. In J. Calderhead (Ed.), *Teachers' professional learning* (pp. 35-50). New York: Falmer Press.

Korthagen, F.A.J., & Wubbels, T. (1991, April). *Characteristics of reflective practitioners: Towards an operationalization of the concept of reflection*. Paper presented at the annual meeting of the American Educational Research Association, Chicago.

LaBoskey, V.K. (1989). An exploration of the nature and stability of reflectivity in preservice teachers. *Dissertation Abstracts International, 51*, 140A. (University Microfilms No. 90-11, 532)

LaBoskey, V.K. (1992). Case investigations: Preservice teacher research as an aid to reflection. In J.H. Shulman (Ed.), *Case methods in teacher education* (pp. 175-193). New York: Teachers College Press.

LaBoskey, V.K. (1993). A conceptual framework for reflection in preservice teacher education. In J. Calderhead & P. Gates (Eds.), *Conceptualizing reflection in teacher development* (pp. 23-38). Washington, D.C.: Falmer Press.

Leinhardt, G. (1989). Math lessons: A contrast of novice and expert competence. *Journal for Research in Mathematics Education, 20*, 52-75.

Mathematical Association of America. (1991). *A call for change: Recommendations for the mathematical preparation of teachers of mathematics*. Washington, D.C.: Author.

Mathematical Sciences Education Board. (1991). *Counting on you: Actions supporting mathematics teaching standards*. Washington, D.C.: National Academy Press.

McKnight, C.C., Crosswhite, F.J., Dossey, J.A., Kifer, E., Swafford, J.O., Travers, K.J., & Cooney, T.J. (1987). *The underachieving curriculum: Assessing U.S. school mathematics from an international perspective, A national report on the Second International Mathematics Study (SIMS)*. Champaign, IL: Stipes.

Miles, M.B., & Huberman, A.M. (1984). *Qualitative data analysis: A sourcebook of new methods*. Beverly Hills: Sage.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.

National Council of Teachers of Mathematics. (1991). *Professional standards for teaching mathematics*. Reston, VA: Author.

National Research Council. (1989). *Everybody counts*. Washington, D.C.: National Academy Press.

Pultorak, E.G. (1993). Facilitating reflective thought in novice teachers. *Journal of Teacher Education, 44*, 288-295.

Replee Proprietary Limited. (1993). *NUD•IST: Non-numerical unstructured data indexing searching and theorizing* [Computer Program]. Melbourne, Australia: Author.

Richert, A.E. (1990). Case methods and teacher education: Using cases to teach teacher reflection. In R. Tabachnick & K. Zeichner (Eds.), *Issues and practices in inquiry-oriented teacher education* (pp. 130-150). London: Falmer Press.

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

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

Siegel, S., & Castellan, N. J., Jr. (1988). *Nonparametric statistics for the behavioral sciences* (2nd ed.). New York: McGraw-Hill.

Spiro, R.J., Coulson, R.L., Feltovich, P.J., & Anderson, D.K. (1988). *Cognitive flexibility theory: Advanced knowledge acquisition in ill-structured domains* (Technical Report No. 441). Champaign, IL: University of Illinois at Urbana-Champaign.

Spiro, R.J., Feltovich, P.J., Coulson, R.L., & Anderson, D.K. (1989). Multiple analogies for complex concepts: Antidotes for analogy-induced misconception in advanced knowledge acquisition. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 498-531). Cambridge, MA: Cambridge University Press.

Spiro, R.J., Vispoel, W.P., Schmitz, J.G., Samarapungavan, A., & Boerger, A.E. (1987). Knowledge acquisition for application: Cognitive flexibility and transfer in complex content domains. In B.K. Britton & S.M. Glynn (Eds.), *Executive control processes in reading* (pp. 177-199). Hillsdale, NJ: Lawrence Erlbaum.

SPSS, Incorporated. (1990). *SPSS for the Macintosh* [Computer Program]. Chicago, IL: Author.

Strauss, A., & Corbin, J. (1990). *Basics of qualitative research*. Newbury Park, CA: Sage.

Teitelbaum, K., & Britzman, D.P. (1991). Reading and doing ethnography: Teacher education and reflective practice. In B.R. Tabachnick & K.M. Zeichner (Eds.), *Issues and practices in inquiry-oriented teacher education* (pp. 166-185). New York: Falmer Press.

Trumbull, D., & Slack, M.J. (1991). Learning to ask, listen, and analyze: Using structured interviewing assignments to develop reflection in preservice science teachers. *International Journal of Science Education, 13*, 129-142.

Valli, L. (1993). Reflective teacher education programs: An analysis of case studies. In J. Calderhead & P. Gates (Eds.), *Conceptualizing reflection in teacher development* (pp. 11-22). Washington, D.C.: Falmer Press.

Van Zoest, L.R. (1993). *Preservice secondary school mathematics teachers' observations of mathematics instruction*. Unpublished manuscript, Illinois State University, Normal, IL.

Van Zoest, L.R., Jones, G.A., & Thornton, C.A. (1994). Beliefs about mathematics teaching held by preservice teachers involved in a first grade mentorship program. *Mathematics Education Research Journal, 6*, 37-55.

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. (1990, November). *Educational and social commitments in reflective teacher education programs*. Paper presented at the annual meeting of the National Forum of the Association of Independent Liberal Arts Colleges for Teacher Education, Milwaukee, WI.

Zeichner, K.M., & Liston, D.P. (1987). Teaching student teachers to reflect. *Harvard Educational Review, 57*(1), 23-48.

Zeichner, K.M., & Tabachnick, B.R. (1991). Reflections on reflective teaching. In B.R. Tabachnick & K.M. Zeichner (Eds.), *Issues and practices in inquiry-oriented teacher education* (pp. 1-21). New York: Falmer Press.

Table 1

Participant Background Information

Group	Name	Sex	Year	Trad	Tran	C&I	Math	UTA	Tutor	GPA	INT
Obser	Dora	F	Jr.	Yes	Yes	216	7	Yes	Yes	3.86	107
	Sara	F	Sr.	No	Yes	215	10	No	No	2.20	109
	Tom	M	Jr.	Yes	Yes	215	7	No	No	3.14	100
	Val	F	Jr.	No	Yes	216	9	No	No	3.44	116
Disc	Bob	M	Jr.	Yes	Yes	216	6	No	No	2.67	136
	Erin	F	Sr.	Yes	Yes	216	8	Yes	Yes	2.50	103
	Fran	F	Jr.	Yes	Yes	215	7	No	No	3.57	110
	Nan	F	Sr.	Yes	No	216	10	Yes	Yes	2.60	117
Facil	Gina	F	Jr.	Yes	Yes	216	6	No	No	2.83	107
	Lynn	F	Sr.	Yes	Yes	216	6	No	No	3.17	106
	Mary	F	Sr.	Yes	Yes	216	7	No	No	3.43	117
	Rand	M	Sr.	Yes	Yes	216	9	No	Yes	3.44	96

Note.

KEY

C&I: Current Department of Curriculum & Instruction course

Disc: Observation-with-discussion treatment group

Facil: Observation-with-facilitated-discussion treatment group

GPA: Mathematics grade point average

INT: Sum of scores on the internal scales of the I.E.O.-Test

Math: Number of mathematics content courses taken

Name: Researcher-generated pseudonyms

Obser: Observation-only treatment group

Trad: Traditional student

Tran: Transfer student

Tutor: Tutor in the University's mathematics assistance laboratory

UTA: Undergraduate teaching assistant for college algebra course

Table 2

Group Means and Standard Deviations for Number of Thoughts, Observations, and Reflections

	Thoughts		Observations		Reflections	
	Pre	Post	Pre	Post	Pre	Post
Observation Only (n = 4)						
Mean	31.0	24.0	19.0	16.8	12.3	7.5
Std Dev	10.8	2.4	7.6	1.9	3.5	0.9
Discussion (n = 4)						
Mean	27.3	32.3	16.0	17.3	12.0	15.3
Std Dev	6.2	2.5	6.2	3.7	1.9	3.0
Facilitated Discussion (n = 4)						
Mean	23.3	25.8	13.5	13.8	9.8	12.0
Std Dev	6.8	6.2	6.1	3.7	1.5	2.5
All Participants (N = 12)						
Mean	27.2	27.3	16.2	15.9	11.3	11.6
Std Dev	8.8	5.4	7.0	3.6	2.7	3.9

Table 3

Mean Ranks for Number of Thoughts, Observations, and Reflections

	Thoughts		Observations		Reflections	
	Pre	Post	Pre	Post	Pre	Post
Observation Only (n = 4)						
Mean Rank	7.50	4.25	7.88	7.13	7.13	2.50
Discussion (n = 4)						
Mean Rank	6.75	9.75	6.25	7.75	8.00	9.63
Facilitated Discussion (n = 4)						
Mean Rank	5.25	5.50	5.38	4.63	4.38	7.38

Table 4

Chi-Squared Approximations for Kruskal-Wallis Test on Post-Treatment Number of Thoughts, Observations, and Reflections (N = 12)

	χ^2_a	df	p
Number of Thoughts	5.151*	2	.0761
Number of Observations	1.719	2	.4234
Number of Reflections	8.309*	2	.0157

[a] The values have been corrected for ties.
* $p < .10$

Table 5

Summary of Reflection Types

	No Justification						Justification by			
	General		Specific		Personal		Experience		Logical Reasoning	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Observation (n = 4)										
Mean	3.5	2.0	5.8	2.8	0.0	0.0	0.5	0.3	2.5	2.5
Std Dev	2.1	1.4	2.9	0.8	0.0	0.0	0.5	0.4	1.5	0.5
Discussion (n = 4)										
Mean	3.0	5.8	3.5	4.0	1.0	1.5	1.0	1.8	3.5	2.3
Std Dev	1.6	3.3	1.1	2.4	0.7	2.6	1.0	1.9	1.1	1.5
Facilitated (n = 4)										
Mean	3.0	5.0	2.0	3.5	1.3	0.0	1.3	0.5	2.5	3.0
Std Dev	1.2	1.6	1.2	0.9	0.8	0.0	1.1	0.9	1.5	1.9
Total (N = 12)										
Mean	3.2	4.3	3.8	3.4	0.8	0.5	0.9	0.8	2.8	2.6
Std Dev	1.7	2.8	2.5	1.7	0.8	1.7	1.0	1.4	1.5	1.4

Table 6

Participants' Abilities to Recall Basic Facts About the Videotaped Lessons and their Observation Experience Preferences

Group	Name	Recall Strength*	Preference
Observation-Only	Dora	Good	Observe
	Sara	Poor	Discuss
	Tom	Poor	Discuss
	Val	Good	Undecided
Observation-with-Discussion	Bob	Good	Discuss
	Erin	Good	Discuss
	Fran	Poor	Observe
	Nan	Good	Discuss
Observation-with-Facilitated Discussion	Gina	Good	Discuss
	Lynn	Good	Discuss
	Mary	Good	Discuss
	Rand	Good	Undecided

* Good: able to remember basic facts about the six videos

Poor: unable to remember basic facts about the six videos

Table 7

Responses to Post-Treatment Interview Question 1: What about the experiences you've had in this study most stands out in your mind?

Group	Name	Initial Response
Observation-Only	Dora	Well, I think it's important to have good rapport with your students, to be on good terms with them, have human communication not just "I'm the teacher and you will do what I say."
	Sara	Well, more than anything when I was looking at the tapes or the classroom, I was really looking for technique and how they presented the material.
	Tom	Well, I'd have to say, I don't remember the guy's name, with the probability, I just really liked that one class a lot. That's the one thing that really sticks out in my mind.
	Val	Differences between teacher styles. Some were more teacher-oriented and others were more student-oriented.
Observation-with-Discussion	Bob	The best thing to me would be the talking among the other people. To me it helped develop a sense of the way other people look at things as opposed to how I look at things.
	Erin	Probably the disagreements we had in the discussion groups.
	Fran	Well, I know this isn't very nice to say, but it seemed like the other people were a lot more critical of the teachers than I was and I guess it - it's kind of strange because they're going to be teachers and they were just so critical and I don't know, they just weren't very nice.
	Nan	The big thing was being able to discuss things with the other members of my group and realizing that I wasn't the only one seeing these things and I wasn't the only one with this impression.
Observation-with-Facilitated-Discussion	Gina	I guess observing things that will help me to become a better teacher. It also helped me to see what the individuals in my group thought was important.
	Lynn	Probably just the difference of teaching techniques that there are. I didn't realize that there was such a variety in how people teach.
	Mary	Probably that there's - I just saw many different - well, not many, just a few different teachers and they all had a bunch of different ways of teaching, styles.
	Rand	I liked the variety of classes, seeing a variety of teachers.

Post-Treatment Interview Protocol

1. What about the experiences you've had in this study stands out most in your mind?

2 a. Which of the five teachers in the videos would you most like to be like?
b. When you first start teaching which one do you think you will most be like?

3. From your experiences in this study, what do you think you've learned about:
a. mathematics;
b. what you want or don't want to do when you teach?

*4. *a. What do you think you learned from interacting with your group members?*
b. What is the most important thing that you learned from the interaction?

5. I would like you rank the five lessons in order of how effective you thought they were, and explain your reasons.

6. In the first interview, when I asked what you thought were key characteristics of a "good" teacher you listed _____. I'm wondering if you would like to make any additions or changes?

7. Another thing that we talked about in the first interview was what you hoped to learn from observing secondary mathematics classrooms. You had said _____. Do you think that happened through watching the videos?

8. If you had a choice when you signed up for your education course between watching 8 videotaped mathematics classroom lessons or watching 4 and then discussing each one for 50 minutes immediately afterwards in a group of four people, which do you think you would prefer? Why?

9. I noticed that in your written reflections, _____.

10. Is there anything else you would like to add?

**asked only of participants in the discussion groups*

Figure 1. Post-Treatment Interview Protocol.

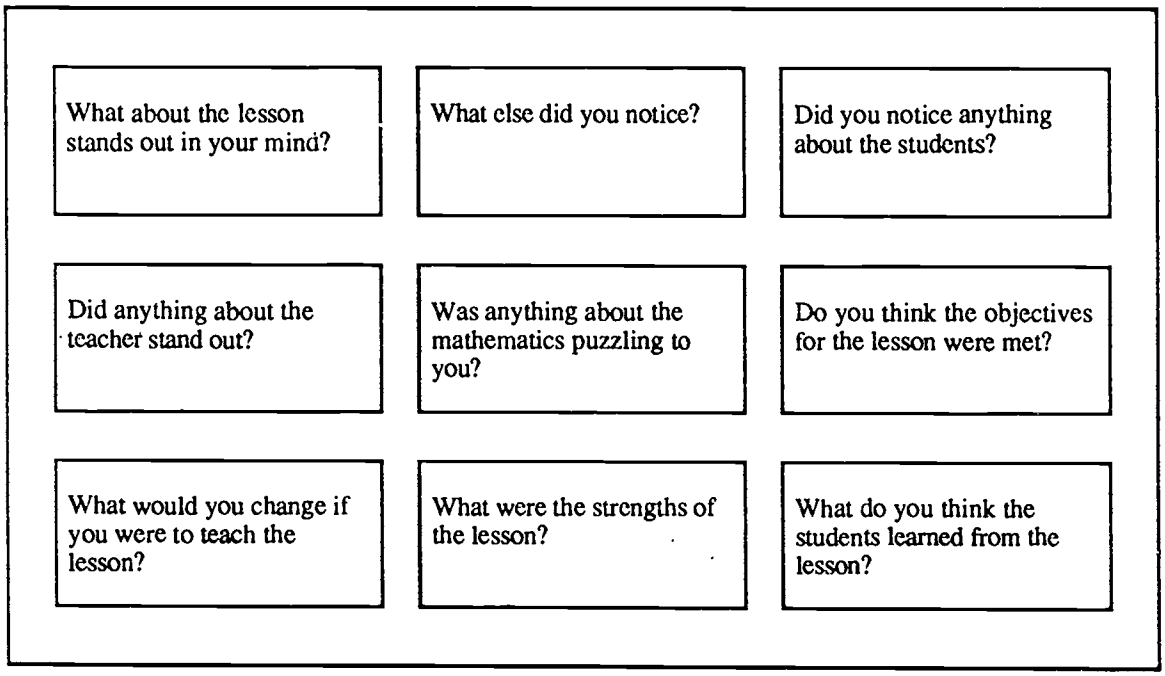

Figure 2. Discussion Questions.

	Low Robustness **general reflection with no justification**	- - - - - - - - - - **specific reflection with no justification**	- - - - - - - - - - **reflection justified by experience**	High robustness **reflection justified by logical reasoning**
Negative	"should have enforced some sort of discipline"	"should have separated the students that were talking"	"that's what they did at my high school and it always seemed to work"	"it will disorient the disruptive students and take awhile for them to be comfortable enough to cause problems"
Positive	"good rapport with the students"	"she seemed to treat the students as responsible adults"	"I always liked teachers who treated me with respect"	"if the teacher treats the students with respect they are likely to reciprocate"
Neutral	"probably thought it was important to check homework first"	"might have wanted to see if the students understood before going on"	"that's what my teachers did and it seemed to make the lesson go smoother"	"it's important to make sure students could do their homework before starting a new topic"

Figure 3. Reflection Rubric with Examples.

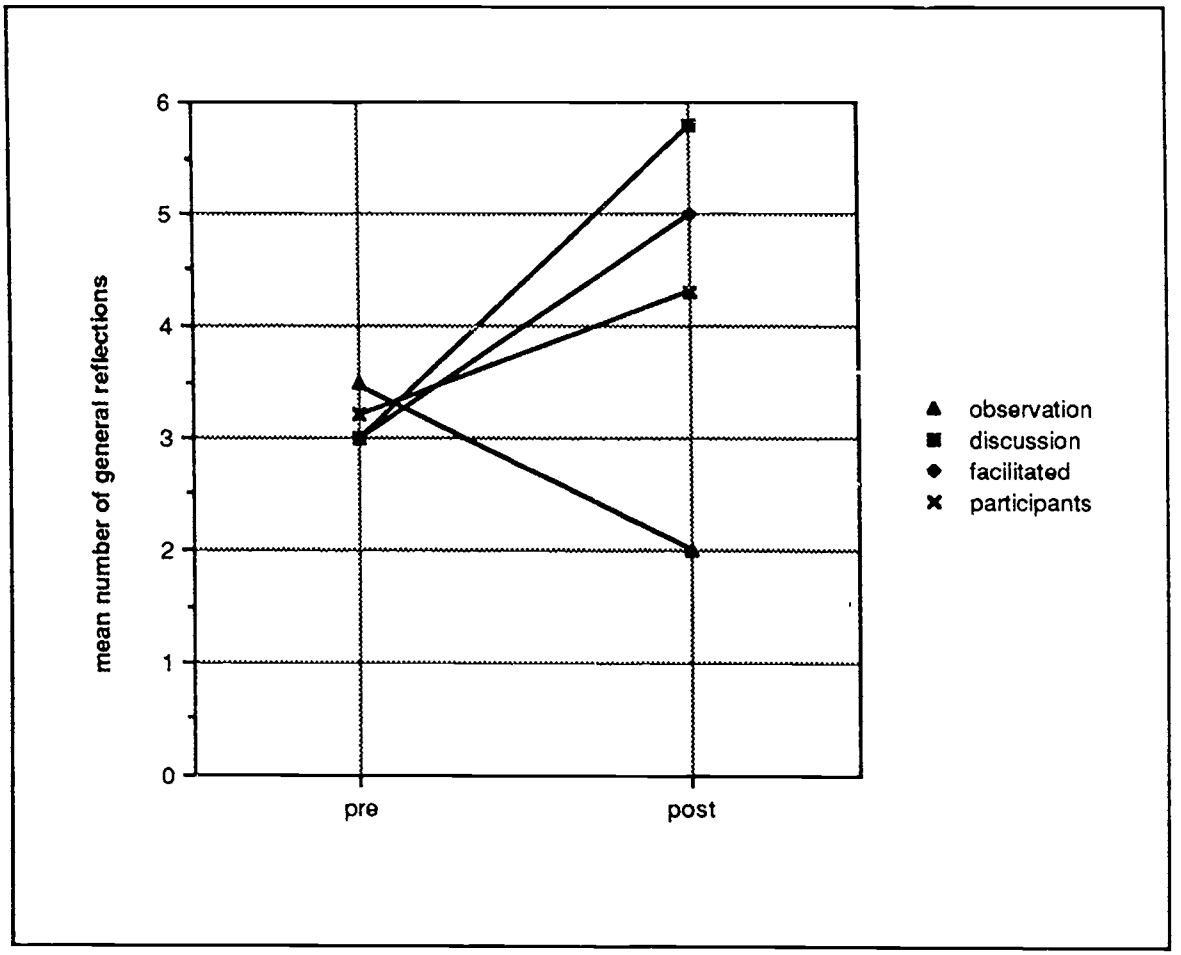

Figure 4. Relationship of Pre- and Post-Treatment General Reflections without Justification.

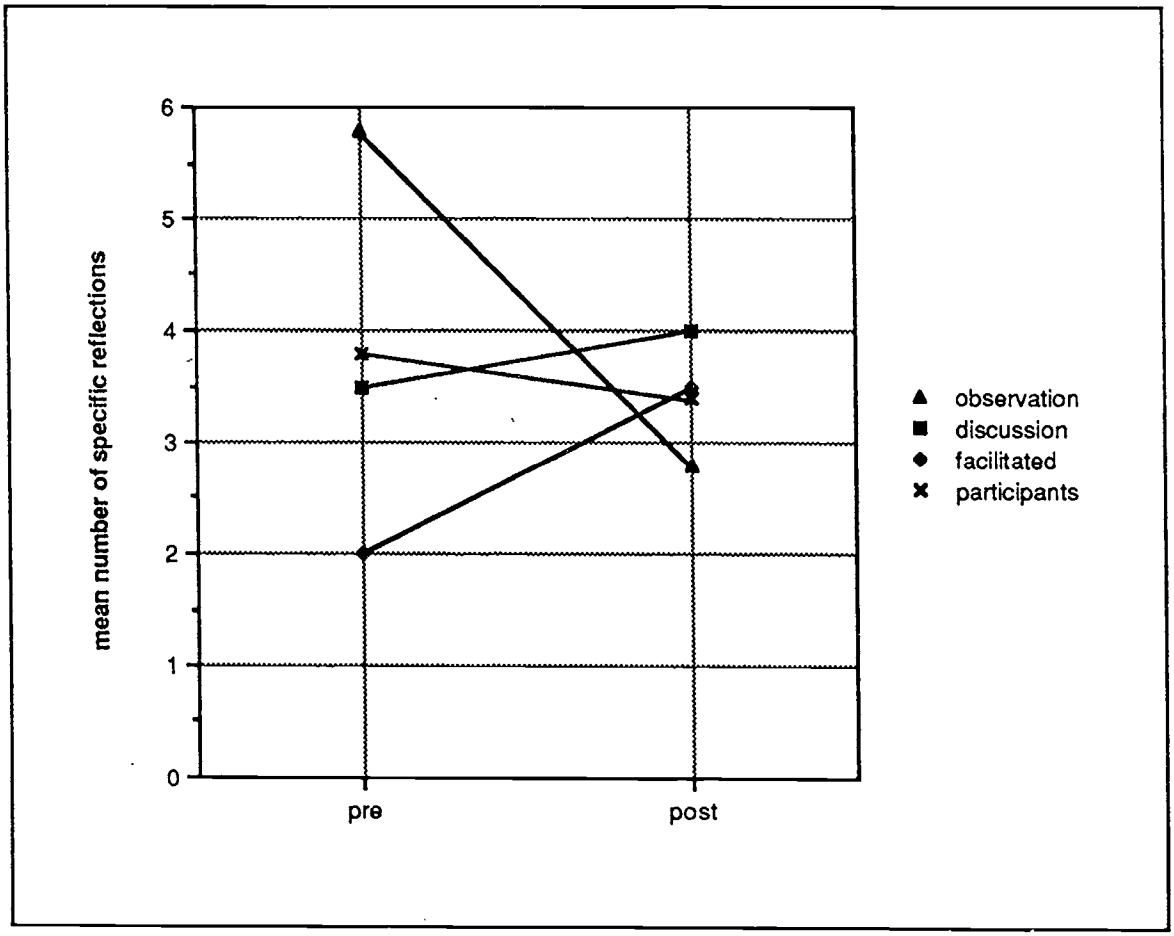

Figure 5. Relationship of Pre- and Post-Treatment Specific Reflections without Justification.

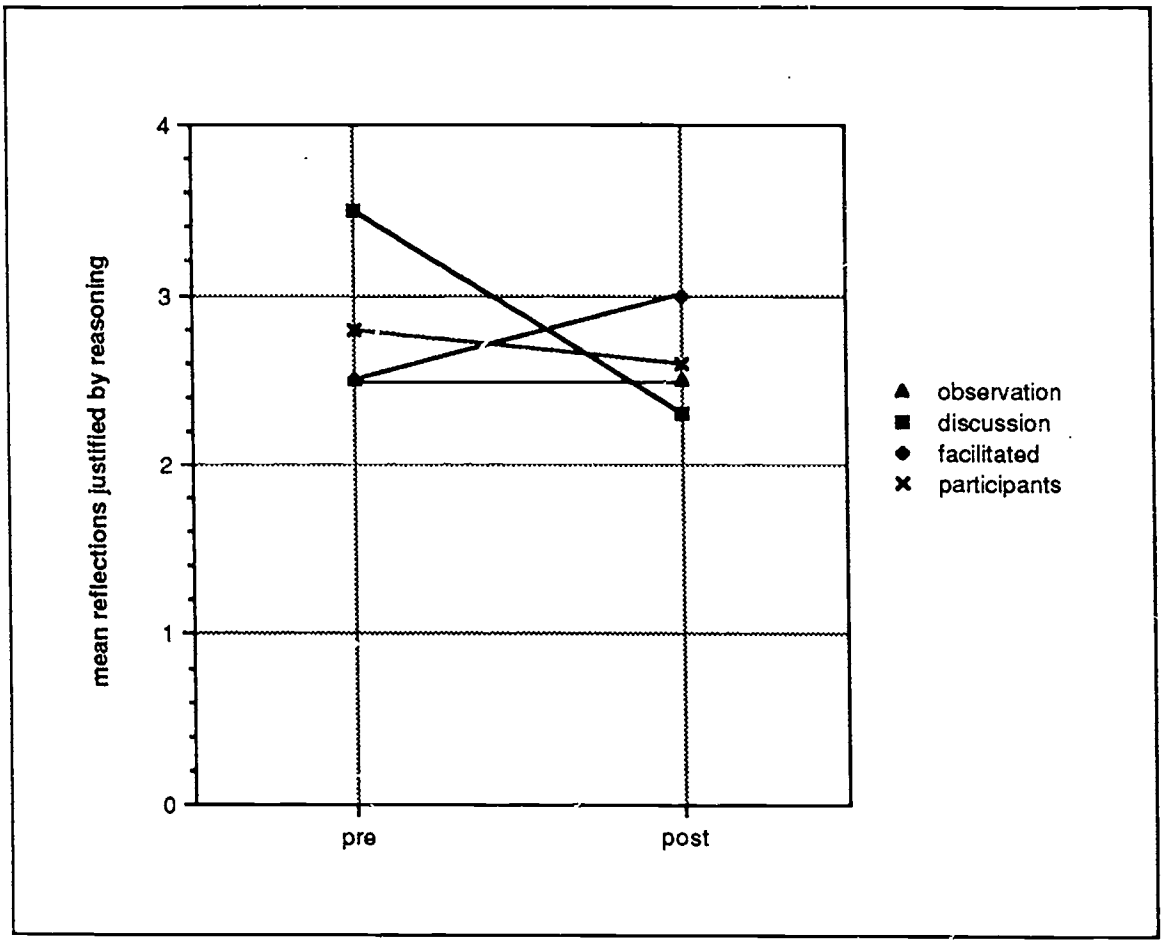

Figure 6. Relationship of Pre- and Post-Treatment Reflections Justified by Logical Reasoning.