Using Self Evaluation to Improve Student Teacher Interns’ Use of Specific Praise

Cassandra L. Keller
Lynn University

Michael P. Brady
and Ronald L. Taylor
Florida Atlantic University

Abstract: This study investigated effects of an audiotape self evaluation intervention on the instructional behavior of three student teacher interns in classrooms for students with various developmental disabilities. A multiple baseline design was used to evaluate of the intervention. Results indicated a positive effect on all interns’ use of specific social praise. Generalization probes indicated two of three interns increased their use of specific social praise in non-targeted content areas. Maintenance probes indicated all interns’ average use of specific social praise was well above baseline. Implications for teacher educators and classroom teachers of students with developmental disabilities are discussed.

Self evaluation is a process in which one makes judgments about the adequacy and effectiveness of performance for the purpose of self improvement. It is the most common form of evaluation used by teachers to improve practice (Airasian, Gullickson, Hahn, & Farland, 1995). The function of self evaluation is to help teachers identify and make decisions about the strengths and weaknesses of their practice, with the intent of improving it (Airasian & Gullickson, 1997). Self evaluation has been identified as a method for promoting increased use of effective practices in teachers (Bullard, 1998; Kilbourn, 1991; Stronge, 1997; Sutherland & Wehby, 2001).

Although research has indicated that data-based self evaluation procedures have been extremely effective in helping practicing teachers increase their use of effective teaching behaviors, very few self evaluation studies have been conducted with student teacher interns. Of the small number of self evaluation studies that have been conducted with teacher interns, most have used rating scales or questionnaires to assess or promote reflections or general perceptions of their teaching (Griffin & Kilgore, 1995; Wheeler & Knoop, 1982). Many educators (Buchanan & Jackson, 1998; Freiberg & Waxman, 1990; Greis, 1986; Kagan & Tippens, 1991; Speer & Zoellick, 1974) have discussed the value of using self evaluation techniques, and how interns might benefit from a more data-based approach to evaluating behavior. This includes attention to accuracy in collecting specific data on one’s own teaching, as well as using these data to make decisions. Frequently, initial attempts at self evaluation do not correspond with more objective, external evaluations (Dougall & Brady, 1998). For example, Briggs, Richardson, and Sefzik (1986) and Wheeler and Knoop found that interns’ self evaluations of their performance tended to be more positive than their supervisors’ ratings. This suggests that interns may not be accurate in self evaluating their performance without specific data on their own instructional behavior.

Specific Praise as an Effective Teaching Behavior

Praise is an instructional behavior that has been shown to be a critical component of effective instruction (Emmer, 1988; Englert, 1984; Hancock, 2000; Sutherland, Wehby, & Yoder, 2002). Yet, studies consistently report that teachers provide an extremely low rate of praise to students in general education and special education classes (Brady & Taylor, 1989; Gable, Hendrickson, Young, Shores & Stowitschek, 1983; Gunter & Coutinho, 1997). Teacher verbal praise has been reported to occur at the rate of less than
five times per hour per student in general education classrooms and less than two times per hour per student in classrooms for students with disabilities (Brophy, 1981; Van Acker, Grant, Van & Henry, 1996; Wehby, Symons, & Shores, 1995). This indicates that students placed in special education classrooms often receive less praise than students in general education classrooms.

Numerous studies indicate teacher praise increases student motivation, task engagement, and actual learning. Sutherland, Wehby, and Copeland (2000) indicated that students’ task engagement increased when teacher praise increased, and decreased when praise was decreased. They found that as teachers increased their use of praise, accuracy of student academic responding improved, resulting in higher academic achievement. Hancock (2000) found that students exposed to verbal praise persisted longer at tasks than a control group that received no praise. He also found it to be a practical way of affecting students’ motivation to engage in behaviors associated with learning.

In a review of 30 years of empirical research on praise and reprimands, Beaman and Wheldall (2000) determined the typical rates used by teachers in elementary through secondary schools. They found students’ academic behavior is praised more often than social behavior and teachers respond more frequently to inappropriate social behavior than to appropriate social behavior. They suggested that teachers frequently fail to take advantage of praise to manage the social behavior of their students.

These studies and others make a strong case for teachers’ use of frequent, specific praise to increase students’ involvement in lessons, their motivation to learn, and their academic performance. Thus, research is needed on techniques that may promote and maintain increased use of effective teaching behaviors such as specific praise. Given the impact student teaching may have on future behavior (Guyton & McIntyre, 1990), it is clear that studying ways to promote interns’ use of effective teaching behavior such as praise is important.

This study examined effects of a self evaluation intervention using audiotape samples of interns’ instructional behavior on the frequency of their (a) specific social praise in classrooms for students with disabilities, (b) specific social praise in non-targeted content areas, and (c) specific social praise after the intervention ended.

Method

Participants and Setting

Three female undergraduate students majoring in Exceptional Student Education at a regional university in South Florida participated. All were completing the student teaching portion of their programs in elementary school settings. They were also enrolled in a special program called Accelerated Induction into Teaching (AIT) in which they served full-time as substitute teachers while completing student teaching requirements. Each intern was assigned a master mentor for 2-3 hours per day for support, and a university supervisor who evaluated the intern’s performance. All interns taught in special education self-contained classrooms with students in grades K-5 with various developmental disabilities.

Interns were European Americans who ranged in age from 24-36 years and had similar prior knowledge and experience with persons with disabilities. Intern 1 (Tracey) had mentored youth with disabilities. Intern 2 (Candice) had worked in a group home for teenagers with disabilities and had a brother with disabilities. Intern 3 (Susan) was a teaching assistant in a classroom with students with disabilities.

Experimental Design

A multiple baseline design across subjects (Baer, Wolfe, & Risley, 1968) was used to evaluate the intervention. The target behavior was frequency of specific social praise. Specific social praise was defined as verbal statements demonstrating approval of a student’s social response that specified the behavior being praised. The treatment was the self evaluation procedure which consisted of (a) a one-time scripted training session with each intern, and (b) the intern’s participation in daily data-collection and graphing activities.

Procedure

To allow interns and students to adjust to one another, no experimental procedures were implemented during the first two weeks of
student teaching. During the 3rd week, interns met with the researcher to receive the equipment for the study. This included; a micro audiotape recorder, a lapel microphone, a box for tapes, and 10 blank pre-dated audiotapes. Interns were then asked which teacher-led content area they wished to target for the study. Teacher-led instruction was defined as instruction to either a small or large group of students that consisted of at least 15 minutes of interaction between intern and students. Teacher-led instruction did not include seatwork or one-to-one instruction. Tracey targeted the content area she called “calendar time.” Calendar time was a time when her students orally practiced functional math and oral language skills. Candice targeted reading and Susan targeted math.

Interns were asked to record their instruction during the targeted content area the following week, and to continue to record the same targeted content area each day until the study was over. During the first three days, no data were evaluated by the interns or researcher. This allowed interns and students to acclimate to the audiotape recorder and data collection procedure (turning recorder on and off, etc.). It also reduced the likelihood that data collection would have a reactive effect on either interns or students (Richards, Taylor, Ramasamy, & Richards, 1999).

During baseline, each intern was asked not to listen to the audiotapes. Tapes were collected regularly (daily or every other day) to assure the request. During intervention, interns did have access to their data tapes, consistent with intervention procedures described under Intervention. When the intervention was removed to assess maintenance, interns were asked once again to refrain from listening to their data tapes, and tapes were again collected.

To ensure random data collection, time intervals of 5 min each were written on index cards. Time intervals started at 0-5 min, then 1-6 min and continued up to 10-15 min. Cards were placed into a bag and drawn each day to determine the interval for data collection.

Once the segment of an intern’s lesson was determined, the researcher recorded frequency of the target behavior since the time of each lesson (5 min) was consistent throughout the study. During baseline, data were collected for at least five days and until at least one intern showed a stable trend. Baseline was considered stable if it did not vary more than 50% from its mean (Alberto & Troutman, 2003). Once baseline became stable, intervention began with the intern with the lowest mean frequency of specific social praise during baseline. The intern with the second lowest mean began intervention next, and so on.

**Intervention**

Intervention for each intern started with an individual training session, including a 60-minute scripted lesson that summarized the nine steps for self evaluation adapted from Airasian and Gullickson (1997) and Sutherland (2000):

1. The intern was asked to make a prediction about her frequency of specific social praise in a 5-min period of teacher-led instruction. According to Airasian and Gullickson (1997), self evaluation is most effective when teachers make predictions about their performance before the data are shared with them.
2. The researcher shared with each intern the average frequency of specific social praise statements per 5 min of instruction from baseline audiotapes.
3. The intern was asked to target the frequency of specific social praise statements for change.
4. The intern was provided with characteristics of effective specific praise and the potential benefits of using effective praise in the classroom.
5. The intern was instructed how to use the recorder and protocol for recording frequency of social praise statements.
6. A training audiotape created by the researcher was then coded by both the intern and the researcher. Results were compared and 100% agreement was obtained with each intern on frequency of specific social praise statements.
7. The intern was asked to set a personal goal for frequency of specific social praise per 5 min of teacher-led instruction. The researcher and intern then brainstormed strategies for use during instruction to help meet the goal. The intern was asked to use the strategy or strategies each day to help reach the personal goal.
8. The intern was asked to record the targeted content area during teacher-led instruction and listen to a 5 min sample of her instructional behavior daily. Gunter and Reed (1996) suggested that a five-minute audiotape sample of instruction was sufficient in providing insight into instructional behavior of teachers.

9. A graph was prepared for each intern. Graphs contained the intern’s baseline data and a goal statement to serve as a reminder of the interns’ target goal. The intern was asked to graph her frequency of specific social praise statements each day until the intervention ended.

A cue card (a visual reminder) was provided to each intern to give her a visual prompt to use effective specific praise frequently. Interns were asked to display the cue cards somewhere in the classroom where they were able to view them while providing instruction.

Generalization and Maintenance

Mentors recorded each intern’s instruction during non-targeted content areas to assess generalization of praise to lessons that had not been targeted. Non-targeted content areas were selected by the mentor when the intern provided teacher-led instruction. These audiotapes were analyzed using the same procedures described earlier. Frequency of praise statements indicated whether generalization to non-targeted content areas occurred.

Mentors also recorded each intern’s instructional lessons during the targeted content area 3½ weeks after the self evaluation intervention ceased. Frequency of each intern’s use of specific social praise statements indicated whether interns maintained improvements over time.

Interobserver Agreement

The researcher and a trained graduate student in special education independently recorded frequency of praise statements in 21% of the taped lessons randomly selected across baseline, intervention, generalization and maintenance. The graduate student was naïve to which condition the tape involved. Percent agreement was calculated by dividing the smaller number of praise statements by the larger number and multiplying by 100. Mean agreement across all conditions for Tracey, Candice, and Susan was 94%, 96%, and 93% respectively.

Results

As indicated in Figure 1, all interns’ praise statements increased when the intervention was introduced.

Tracey. Frequency of specific social praise increased from baseline ($M = 0.3$, range 0-1) to intervention ($M = 4.8$, range 1-13) for Tracey. A Split-middle procedure (White & Haring, 1980) was used to determine the trend lines. Data indicated a rapidly increasing, but variable trend during intervention. Mean frequency of specific social praise during maintenance was 4.4 (range 0-10). Although mean frequency of social praise nearly matched the intervention results during maintenance, a decreasing trend was evident.

Candice. Frequency of Candice’s specific praise increased from baseline ($M = 0.1$, range 0-1) to intervention ($M = 2.9$, range 1-4). Data indicated a gradually increasing stable trend during intervention. Mean frequency of specific social praise during maintenance was 2.8 (range 0-6). Although mean frequency remained fairly stable from intervention to maintenance, the trend was gradually decreasing.

Susan. Frequency of Susan’s specific social praise increased from baseline ($M = 0.5$, range 0-4) to intervention ($M = 4.5$, range 3-5). Data indicated a flat trend with high stability, during intervention. During maintenance, mean frequency of specific social praise was 5.8 (range 3-9), which indicated an increase in level from intervention to maintenance. Further inspection of the data during maintenance showed a continually increasing trend as well.

Generalization Probes

Because mentors were not always available during interns’ instruction, generalization probes were not collected equally across all conditions. Indeed, some probes were only administered once during baseline and intervention for each intern. Because of this limitation, means are presented along with number of probes in each condition in Table 1.

Tracey. Frequency of Tracey’s specific social
Figure 1. Specific social praise used during teacher-led instruction with trend lines.
praise slightly increased from baseline (M = 0, range 0) to intervention (M = 0.7, range 0-2) during non-targeted contents areas (see Figure 2). However, only one probe was administered during baseline because the mentor was not present during this short baseline to collect data during instruction. Mean frequency of specific social praise during maintenance was 1.4, (range 0-3). Although the level remained fairly stable from intervention to maintenance, the trend during both phases was decreasing.

Candice. Frequency of Candice’s specific social praise decreased from baseline (M = 0.8, range 0-2) to intervention (M = 0.0, range 0) during non-targeted content areas. Only one probe was administered during intervention because the mentor was not present when Candice provided teacher-led instruction during this phase. Mean frequency of social praise during maintenance was 4.4, (range 0-12) and indicated a rapidly increasing, yet variable trend.

Susan. Frequency of Susan’s specific social praise increased from baseline (M = 0.6, range 0-3) to intervention (M = 2.5, range 0-7) during non-targeted content areas and indicated a rapidly increasing, stable trend during intervention. Mean frequency of social praise during maintenance was 5.5, (range 2-12) and indicated an increasing yet variable trend.

Discussion

The purpose of this study was to examine effects of a self evaluation intervention on the instructional behavior of student teacher interns. All three interns immediately increased their use of the targeted behavior (specific social praise) when the intervention was implemented. Social praise use for two interns showed an increasing trend during intervention and the third intern showed stable and substantial increases. Interns also maintained their use of social praise throughout the study.

These findings are consistent with other studies that involved data-based self evaluation procedures to help teachers increase use of effective teaching behaviors (Sutherland & Wehby, 2001; Wright, 1998). These researchers asked classroom teachers to use self evaluation procedures to increase their use of praise statements. Both studies indicated that self evaluation procedures were successful at increasing teachers’ use of praise statements.

An additional question was whether any increases would generalize to content areas other than the one targeted. When examining the generalization probes across conditions, two of the three interns did increase their use of specific social praise during non-targeted content areas. The third intern, however, decreased her use of social praise statements during non-targeted content areas. Although results for this question are generally positive they are also inconclusive due to the paucity of data during these conditions. Thus, results of generalization probes must be considered with caution.

Probes during non-targeted content areas in the maintenance condition were administered more consistently and indicated in-

<table>
<thead>
<tr>
<th>Intern</th>
<th>Condition</th>
<th>Mean Frequency of Social Praise (Number of Probes)</th>
<th>Content Areas Probed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tracey</td>
<td>Baseline</td>
<td>0.0 (1)</td>
<td>Language</td>
</tr>
<tr>
<td></td>
<td>Intervention</td>
<td>0.7 (7)</td>
<td>Reading, Math, Writing, Following Directions Social Studies</td>
</tr>
<tr>
<td></td>
<td>Maintenance</td>
<td>1.4 (5)</td>
<td>Social Studies, Math, Reading</td>
</tr>
<tr>
<td>Candice</td>
<td>Baseline</td>
<td>0.8 (4)</td>
<td>Writing</td>
</tr>
<tr>
<td></td>
<td>Intervention</td>
<td>0.0 (1)</td>
<td>Math</td>
</tr>
<tr>
<td></td>
<td>Maintenance</td>
<td>4.4 (5)</td>
<td>Science, Spelling, Writing, Math</td>
</tr>
<tr>
<td>Susan</td>
<td>Baseline</td>
<td>0.6 (7)</td>
<td>Spelling, Reading</td>
</tr>
<tr>
<td></td>
<td>Intervention</td>
<td>2.5 (4)</td>
<td>Reading, Spelling</td>
</tr>
<tr>
<td></td>
<td>Maintenance</td>
<td>5.5 (6)</td>
<td>Reading, Social Skills</td>
</tr>
</tbody>
</table>
Figure 2. Specific social praise used during non-targeted content areas (generalization) with trend lines.
creased use of social praise statements over baseline.

The final question was to examine whether effects of the intervention would maintain after the intervention was discontinued. All interns’ use of social praise statements during maintenance were above baseline. One intern increased social praise statements made from intervention to maintenance while the other two maintained the level reached during intervention.

Although the intervention was successful in maintaining interns’ use of social praise for 3½ weeks following the cessation of the intervention, it is possible that additional cues or prompts would have helped them continue to use social praise statements at a more robust level. Also the duration of the intervention may have been too short to build durable responding. Other researchers have shown that longer intervention periods and more gradual dismantling of interventions produce more durable responding when self management interventions are removed (Dougal & Brady, 1998). This assists newer behaviors to develop as a regular part of a pattern of interactions. Unfortunately, time constraints associated with student teaching in this study precluded the opportunity for us to study or strengthen that pattern.

This research adds to the body of literature on data-based self evaluation procedures as a valuable means for promoting increased use of effective instructional practices with teachers (Bullard, 1998; Kilbourn, 1991; Sutherland & Wehby, 2001). However, the results from this study also indicated that a data-based self evaluation strategy is effective for increasing student teacher interns’ use of effective practices. These results hold important implications for teacher educators, classroom teachers and teachers in training. For teacher educators, the study found that the self evaluation process was feasible for interns to use to help improve instruction. Teacher educators may want to incorporate a similar strategy into existing student teaching requirements to help interns change their own behavior. This practice could, at a minimum, supplement the external feedback and prompts by supervisors or cooperating teachers used in most teacher education programs. Interns could then learn to examine their own teaching practices, and learn strategies that they can use throughout their teaching careers. As interns become practicing teachers, they may find self evaluation strategies to be effective tools to help them improve their own practice.

In future investigations, researchers should not only look at interns’ instructional behaviors, but how teacher self evaluation interventions might affect student social and academic performance. This could be achieved by determining students’ task engagement and academic products to determine if the intervention affected academic performance. Additionally, researchers could see if other teacher behaviors (besides social praise) change as a result of intervention. For example, teachers may inadvertently alter the frequency with which they provide academic praise, general praise, and/or teacher reprimands. These and other teacher variables play an important role in students’ overall success in school.

References


Self Evaluation / 375
(ERIC Document Reproduction Service No. ED428074)


Received: 10 March 2004
Initial Acceptance: 26 April 2004
Final Acceptance: 15 July 2004