The Effect of Specific Feedback on Critical Reflection of Physical Therapy Students During Internships

Deanna Dye
Karen Wilson Scott
Lee Ostrom
Nancy Devine
Jennifer Leight

Abstract
An essential component of expert professional practice is a practitioner’s ability to critically reflect on one’s performance. The purpose of this study was to determine the effect of specific electronic feedback provided by the coordinator of clinical education on students’ critical reflection ability displayed in weekly journal writings during two consecutive, full-time clinical internships. In a quasi-experimental cross-over cohort design, 15 physical therapy (PT) students were stratified-randomly into two groups. Independent raters scored students’ weekly journal submissions on a 5-point scale. Group one received specific feedback during the first internship; Group two received specific feedback during the second internship. While there was no statistical difference in performance between the two groups, it is undetermined whether specific electronic feedback alters critical reflection ability as reflected in journal writing. It appears that timing and continuing feedback once provided may influence effectiveness. Continued research is warranted.

Introduction
An essential component of expert professional practice is the practitioners’ critically reflection on their performance (Clouder, 2000; Cross, Liles, Conduit, & Price, 2004; Gandy & Jensen, 1992; Jensen, Gwyer, Shepard, & Hack, 2000; Perkins, 1996; Shepard & Hack, 1999). The influence of critical reflection (CR) on expert practice occurs in the social science professions (Bain, Ballantyne, Packer, & Mills, 1999; Campoy & Radcliffe, 2002; Crotty & Allyn, 2001; Fund, Court, & Kramarski, 2002), and in the applied science professions such as nursing, medicine, and PT (Baird, 1996; Burrows, 1995; Clouder, 2000; Jensen et al., 2000; Jensen & Paschal, 2000; Kember, Jones, Loke, McKay, Sinclair, Tse et al., 1999; Shepard & Jensen, 1990). In this study, CR was defined as “the higher order intellectual and affective activities in which physiotherapists engage to critically analyze and evaluate their experiences in order to lead to new understandings and appreciation of the way they think and operate in the clinical setting” (Donaghy & Morss, 2000, p. 6).

Critical reflection in PT appears to be well
established (Clouder, 2000; Cross et al., 2004; Gandy & Jensen, 1992; Jensen et al., 2000; Perkins, 1996; Shepard & Hack, 1999). The impact of specific intrapersonal and extrapersonal factors, and the influence of specific feedback on students’ CR performance has not been established. Multiple CR definitions and operational scales complicate investigation (Bain et al., 1999; Cross, 1993; Donaghy & Morss, 2000), as do various CR assessment methods and research designs. Researchers have assessed CR ability through the use of videotape (Crotty & Allyn 2001, Liimatainen, Poskiparta, Karhila, & Sjogren, 2001), questionnaires (Kember et al., 2000; Sobral, 2000; 2001), guided reflective dialogue (Donaghy & Morss, 2000; Johns, 1996; Plack & Santsier, 2004), journals (Fund et al., 2002; Kember et al., 1999; Plack et al., 2005; Williams et al., 2000; Williams et al., 2002; Wong et al., 1995), and portfolios (Routledge & Willson, 1997).

Numerous studies endorse journaling as promoting reflection (Bain et al., 1999; Boud, 2001; Fund et al., 2002; Kerka, 1996; Riley-Doucet & Wilson, 1997; Williams et al., 2002; Woodward, 1998). Wong et al. (1995) found writing accurately indicated reflection ability. Writing itself might develop reflection ability (Brookfield, 1995; Hatton & Smith, 1994; Hiemstra, 2001).

In quantitative studies reviewed, coding and analyzing journal writing was consistently used for measuring student reflection levels (Kember et al., 1999; Williams et al., 2000; Wong et al., 1995). Valid measurement depends on a clear operational definition, appropriate for the skill and context. Several studies have operationally defined and measured CR (Bain et al., 1999; Donaghy & Morss, 2000; Fund et al., 2002; Johns & Graham, 1996; Kember et al., 1999; Kember et al., 2000; Liimatainen et al., 2001; Morrison, 1996; Plack & Santsier, 2004; Riley-Doucet & Wilson, 1997; Routledge & Willson, 1997; Williams et al., 2000; Williams et al., 2002; Wong et al., 1995). The reliability of a coding system to rate students’ level of reflection as demonstrated in their journal writing has been investigated (Kember et al., 1999; Williams, 2000; Wong et al., 1995). Donaghy and Morss (2000) and Williams et al. (2000), adopted Boud, Keogh, and Walker’s (1985) 5-level schema for developing reflection in PT professionals.

Many authors suggest that feedback may encourage advanced reflection ability (Bain et al., 1999; Beveridge, 1997; Boud & Walker, 1998; Boud, 1999; Donaghy & Morss, 2000; Johns & Graham, 1996; Kember et al., 1999; Plack & Santsier, 2004; Whipp, 2003; Williams et al., 2000; Williams et al., 2002; Wong et al., 1995). Williams et al. (2000), speculate that low change in students’ CR over time could result from absence of feedback. Some authors theorize that interactive dialogue deepens CR (Ballantyne & Packer, 1995; Hatton & Smith, 1994). Others consider one-directional feedback effective (Bain et al., 1999).

Qualitative studies on expert professional practice observe that CR occurs during a complex or perplexing situation (Jensen et al., 2000; Schön, 1983). Mezirow (1998), Brookfield (1995), and Schön (1983) suggest that a professional should challenge pre-set assumptions. Someone questioning the practitioner’s action can stimulate reflection.

Purpose
The purpose of the reported study was to determine the effect of specific written electronic feedback provided by the coordinating faculty member on PT students’ CR ability displayed in their written weekly journal entries describing their clinical performance during two consecutive full-time clinical education experiences.

Method
Sample
The University of Idaho Human Subjects Committee approved this study. The researcher invited 17 PT students to participate. Fifteen volunteered and provided informed consent. Each participant was enrolled in two consecutive full-time, 8-week clinical education internships (PTOT-734 and PTOT-735), performed at multiple clinical sites with an assigned
licensed PT clinical instructor (CI) to directly supervise and evaluate.

The researcher, who served as the academic clinical coordinator (CC), gathered participant demographics: age, gender, marital status, number in household, and grade point average (GPA). Participants were six males and nine females, aged 21 to 35. Sixteen held bachelors degrees and one a masters. Five were single and ten were married; four lived alone, five lived with one other, and six lived with two or more people. All GPAs were above 3.6.

Procedure

Participants were assigned to the initial feedback group (Group IF) or the later feedback group (Group LF) through a stratified random process considering age and gender. During the internships, students emailed weekly self-SOAP notes (Author, 2005) describing a significant learning event during the week to the CC. Instructions were:

Students should focus on a significant learning event, situation or issue that occurred during the week. Each student will submit a self-SOAP note addressing the event, situation or issue. The note will be written in a SOAP note format and be completed at least weekly. S = subjective perceptions of the event, situation, or issue; O = objective information on the event, situation or issue; A = assessment of the event, situation, or issue; P = plan for future action based on event, situation, or issue.

In addition to submitting self-SOAP notes, the participants provided weekly site factors ratings using a Likert scale (1 = none or not helpful/challenging and 5 = constant or extremely helpful/challenging). The three site factors included: amount of CI feedback, feedback helpfulness, and degree of internship challenge. At each internship conclusion, participants provided an overall internship score, an overall CC feedback helpfulness score, and an estimated average self-SOAP note writing time.

Two independent raters (PT program faculty previously trained in CR rating) evaluated participants' self-SOAP notes content on Williams et al. (2002) 1-5 scale (Table 1). They discussed discrepancies and reached joint decisions.
<table>
<thead>
<tr>
<th>Level</th>
<th>Observed Criteria</th>
<th>Critical question answered</th>
<th>Probing questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>No analysis of the learning event, issue, or situation has occurred</td>
<td>Entry is descriptive in nature. No question is answered.</td>
<td>Ask what occurred? What did you do? What did the patient do?</td>
</tr>
<tr>
<td>1</td>
<td>Describes the learning event, issue, or situation.</td>
<td>What happened?</td>
<td>What is your reaction to the learning event, issue, or situation? Why did it happen?</td>
</tr>
<tr>
<td>2</td>
<td>Analyzes/re-evaluates the learning event, issue, or situation in relation to prior knowledge, feelings, or attitudes.</td>
<td>What is your reaction to the learning event, issue, or situation? Why did it happen?</td>
<td>What is the value of the learning event, issue, or situation that has occurred? Are the new knowledge, feelings, or attitudes. About the learning event, issue, or situation correct?</td>
</tr>
<tr>
<td>3</td>
<td>Verifies/confirms the learning event, issue, or situation in relation to prior knowledge, feelings, or attitudes.</td>
<td>What is the value of the learning event, issue, or situation that has occurred? Are the new knowledge, feelings, or attitudes. About the learning event, issue, or situation correct?</td>
<td>What is your new understanding of the learning event, issue, or situation?</td>
</tr>
<tr>
<td>4</td>
<td>Relates 1, 2, and 3 above to gain a new understanding of the learning event, issue, or situation.</td>
<td>What is your new understanding of the learning event, issue, or situation?</td>
<td>How will you approach the same or similar event, issue, or situation in the future?</td>
</tr>
<tr>
<td>5</td>
<td>Indicates how the new learning event, issue, or situation will affect future behavior. Determines the clarification of an issue, the development of a skill, or the resolution of a problem.</td>
<td>How will you approach the same or similar event, issue, or situation in the future?</td>
<td>No specific questions. Did you hold any assumptions about the situation? Do you hold any assumptions about the future?</td>
</tr>
</tbody>
</table>

Note: Williams, et. al., 2002
### Table 2: Examples of Feedback and General Encouragement

<table>
<thead>
<tr>
<th>Specific Feedback</th>
<th>General Encouragement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 1 – “Try to focus on a single [learning] event with your next note. Then let me know: ‘What did you learn from it? Of what you learned, what is worth retaining?’”</td>
<td>“Thanks for the note. Sometimes reactions can really surprise you.”</td>
</tr>
<tr>
<td>Level 2 – “… ‘In your future SOAP notes try to elaborate on exactly what you learned from the experience you describe and also what you might incorporate into your future practice.’”</td>
<td>“Thanks for the note and ratings. It must be so frustrating not to be able to achieve anything that was desired during the session.”</td>
</tr>
<tr>
<td>Level 3 – “So what do you think contributes to motor function overall? Is the sensory information that critical? What contribution does it seem to make compared to other factors?”</td>
<td>“Thanks for your note. Keep it up.”</td>
</tr>
<tr>
<td>Level 4 – Communication is such a key component. …..Often [parents] also don’t want to engage in out-of-school therapy because of additional costs. So how do you think you would curb these misunderstandings when you are practicing?”</td>
<td>“What an interesting experience! I hope you have a great 8 weeks.”</td>
</tr>
<tr>
<td>Level 5 – “How do you think you’ll be able to gather information more quickly? …. Do you think it is simply the complexity of the patient, or do all similar patients pose the same challenges?”</td>
<td>“Sounds like a crazy week. … I hope it is coming together. See you tomorrow.”</td>
</tr>
</tbody>
</table>

The CC copied submitted self-SOAP notes into one document, removing identifiers. Copying, pasting, and blinding self-SOAP notes averaged 1 hour per week. Raters spent an average of 3 hours per week reading and rating submissions. The raters returned scored notes to the CC, who paired assigned ratings with the cues (Table 1) to design specific feedback (Table 2). Feedback was emailed to students within 4 days of receiving self-SOAP notes. Composing and sending feedback and/or general encouragement averaged 3 hours per week. Receiving, copying, blinding, rating, and responding averaged 28 minutes per student per week.

The CC recorded site ratings and CR scores on a Microsoft Excel? spreadsheet. Student responses to specific questions in subsequent self-SOAP notes indicated feedback recognition. Receipt of feedback was recorded as a dichotomous yes/no. The CC entered and verified data, performing statistical analyses utilized SPSS Version 14.0? software and Microsoft Excel? \((p=0.10)\).  

### Results

#### Data Characteristics

Many weeks students did not submit timely notes, yet, CR ratings remained consistent. An overall average CR score for both completed 8-week courses was representative of the Groups’ performance via statistical analysis. Missing or absent submissions were noted in analysis.
Group Comparisons

Prior to treatment, no apparent or statistical difference was noted between groups on initial CR ratings. For perceived site factors; amount of feedback, helpfulness of feedback, and degree of challenge during PTOT-734, Group LF rated the feedback amount and degree of challenge significantly higher than Group IF. During PTOT-735, Group LF rated feedback amount and feedback helpfulness significantly higher than Group IF ($p=0.02$). There was no significant difference in any site factor ratings within groups between the two internships. There was a significant difference between groups in self-SOAP note writing time (Group IF > Group LF; $p=0.06$), however, three students did not report usable time estimates in Group IF.

Using simple descriptive statistics of the percentage of “yes” notations, Group IF feedback, was overtly recognized 57% of the time; Group LF feedback, 43% of the time. Delayed submissions hampered analysis. Only 4 of the 11 students noticed a change in CC feedback helpfulness. One student rated the CC specific feedback as less helpful than general encouragement (Figure 1). Eleven students’ ratings were analyzed, because two students (in Group IF) did not submit.

PTOT-734 group average CR scores during PTOT-734 were compared to PTOT-735 CR scores. Statistically, there was no difference between groups or within groups over time. Descriptively, Group IF scored higher when receiving feedback compared to not receiving feedback, while Group LF remained unchanged (Figure 2).

Figure 1: Student scores of CC helpfulness when receiving and not receiving feedback

Note: Students rated the CC on a 1-5 scale regarding helpfulness of provided feedback provided. 1 = not at all helpful and 5 = extremely helpful.
Figure 2: Weekly CR scores per group overtime

Note: Weeks 1A-8A = PTOT-734; Week 1B – 8B = PTOT-735; Group 1 = Group IF; Group 2 = Group LF

**Discussion**

No statistical feedback effect was found through inferential analysis. Descriptively, Group IF declined substantially in CR scores when feedback was removed. Whether the students received CC feedback is unknown. Feedback was provided through one-directional emails, which may not have been read prior to the subsequent self-SOAP note submission. Questions within the feedback may not have been specific enough to consistently stimulate student response. Little difference in perceived CC helpfulness was identified. Logistical issues and delays in self-SOAP note submissions may have impacted feedback effectiveness.

Descriptively, Group LF appeared to perform better than Group IF. Group IF demonstrated an apparent drop in scores upon feedback removal during PTOT-735 (Figure 2). Scores gradually improved, but did not reach the PTOT-734 level. Recognizing that the groups were equal (CR ability, age, gender, and GPA), other differences between groups were considered. Clinical sites factors may have influenced students’ performance and feedback timing.

Group IF received feedback immediately, then had it removed; whereas, Group LF received general feedback until the second internship. Group IF CR scores may have declined during PTOT-735 because of specific feedback withdrawal after receiving it. It is unknown whether Group IF would have declined if
specific feedback were maintained. Group LF experienced little CR score change during PTOT-734 compared to PTOT-735. Since Group LF received no CC feedback during the first 8-weeks, students may have disregarded CC feedback. Disregard might have occurred secondary to only general encouragement during the first 8-weeks, or specific feedback may have countered student complacency over time enabling students to maintain prior performance levels.

As neither group demonstrated a significant change in CR ability, specific feedback provided via email might have no effect. Or perhaps, specific feedback and general encouragement assisted in maintaining student performance.

**Suggestions for Further Study**

Further research on CR development of professional students is warranted. This study could be replicated with larger sample sizes. Different designs could be employed to enhance control of extraneous variables. Investigation of different types and timing of feedback should continue. Further studies similar to the one by Bain, et al. (1999) should compare specific and general feedback methods with timing. CR ability and specific clinical performance outcomes could be investigated to understand relationships.

**Conclusion**

Although the results were inconclusive, it appears feedback initiated should continue throughout the project for sustained higher levels of critical reflection performance. Delays in feedback may diminish feedback effectiveness. Specific feedback or general encouragement may have similar effects on student performance. Further research is warranted.

**References**


Deanna Dye is Assistant Professor of Physical Therapy at Idaho State University, Pocatello, ID.

Karen Wilson Scott is Professor of Human Resource Training and Development at Idaho State University, Pocatello, ID.

Lee Ostrom is Professor of Engineering Technology at the University of Idaho, Idaho Falls, ID.

Nancy Devine is Associate Professor of Physical Therapy at Idaho State University, Pocatello, ID.

Jennifer Leight is a Physical Therapist and Clinical Psychologist in Bozeman, MT.