An Investigation of a Tutor Evaluation Scale for Formative Purposes in a Problem-Based Learning Curriculum

John A. Hay

Key Words: education • teaching methods • validity of tests

Problem-based learning (PBL) curricula are of increasing interest to educators within occupational therapy and other allied health professions (Bruhn, 1992; Hay, 1995; Jacobs & Lyons, 1992; Royeen, 1995; Saarinen & Salvatori, 1994; VanLeit, 1995). PBL approaches involve small groups of students who meet with a tutor to examine health care problems that highlight learning needs. With the tutor’s guidance, these students develop learning objectives and establish strategies to meet them. The knowledge gained through this process appears to be well retained because the learning occurs in context and is responsive to student needs (Norman & Schmidt, 1992).

Tutors are often skilled clinicians who are interested in education and are willing and able to become involved in PBL curricula. To provide feedback to tutors about which aspects of their performance need improvement, reliable evaluation tools are required. Teaching evaluation tools commonly used in traditional curricula are inappropriate for tutor evaluation because they measure didactic teaching skills rather than the facilitation and guidance skills required for student-directed learning in small groups (Irby, 1984; Rippey, 1981).

The roles, requisite skills, and attitudes of tutors involved in student-directed learning have been well described (Barrows & Tamblyn, 1980; Knowles, 1975; Maxwell & Wilkerson, 1990; Schmidt & Moust, 1995). These attributes include the ability to facilitate the group process, communicate clearly, enable individual learning objectives, and provide constructive feedback. Each attribute can be evaluated by the student on the basis of their experience with the tutor as well as through direct observation, videotaping, or other more demanding means. The area of tutor evaluation has been largely unaddressed in the literature (Wakefield, Keane, & Hay, 1992). Additionally, recent investigations of tutor effectiveness that rely heavily on student evaluations provide few data on the psychometric properties of the evaluation tools (Schmidt & Moust, 1995).

On the basis of a review of existing measures and discussions with experienced tutors, Hay (1996) and Hay and Schmuck (1993) developed a 19-item scale specifically for evaluation of tutors. Eighteen of the items address specific performance aspects (e.g., the provision of verbal feedback). The final item provides a global rating (see Figure 1). The scale has good interobserver and test–retest reliabilities. Fourteen items had test–retest (over a 5-day interval) reliabilities of more than .70, and one had a reliability of less than .50. Seven items had

John A. Hay, PhD, is Assistant Professor, Faculty of Physical Education and Recreation, Brock University, St. Catharines, Ontario L2S 3A1, Canada. At the time of this study, he was, Assistant Professor, School of Rehabilitation Science, McMaster University, Hamilton, Ontario, Canada.

This article was accepted for publication January 25, 1996.
In order to maintain high educational standards and to facilitate group functioning, the Programmes require feedback from students regarding the performance of faculty. This information will be used in considering reappointments, to guide faculty development, to assist in considerations of promotion and tenure. These evaluations are an integral part of the feedback that tutors receive and will enable tutors to continue their pursuit of excellence by the identification of both strengths and areas in need of improvement.

Use the following scale for your ratings.

<table>
<thead>
<tr>
<th>inadequate</th>
<th>less than adequate</th>
<th>slightly less than adequate</th>
<th>adequate</th>
<th>slightly more than adequate</th>
<th>more than adequate</th>
<th>outstanding</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
</tbody>
</table>

This Tutor:

1. Provided clear insight into the objectives and expectations of the course. 1 2 3 4 5 6 7
2. Discussed students' expectations of the course. 1 2 3 4 5 6 7
3. Assisted students in planning to meet course objectives. 1 2 3 4 5 6 7
4. Encouraged effort and rewarded student contributions. 1 2 3 4 5 6 7
5. Consistently provided reasonable, constructive, verbal feedback. 1 2 3 4 5 6 7
6. Was objective in formal student evaluations. 1 2 3 4 5 6 7
7. Was sensitive to students' learning needs. 1 2 3 4 5 6 7
8. Displayed concern and empathy for students. 1 2 3 4 5 6 7
9. Was enthusiastic about educational role. 1 2 3 4 5 6 7
10. Provided a reasonable degree of student autonomy in establishing learning objectives. 1 2 3 4 5 6 7
11. Provided a reasonable degree of student autonomy in selecting learning strategies. 1 2 3 4 5 6 7
12. Communicated clearly with students. 1 2 3 4 5 6 7
13. Assisted in the development of individuals' communication skills within the group. 1 2 3 4 5 6 7
14. Facilitated development of student problem solving skills. 1 2 3 4 5 6 7
15. Effectively assisted students to develop their reasoning skills. 1 2 3 4 5 6 7
16. Challenged students to develop their knowledge to an appropriate level. 1 2 3 4 5 6 7
17. Stimulated interest in course content. 1 2 3 4 5 6 7
18. Assisted the group in determining when an appropriate level of understanding had been reached. 1 2 3 4 5 6 7

Figure 1. Sample tutor performance evaluation form.

interobserver reliabilities of more than .70, and two had reliabilities less than .50. The reliability of the total evaluation results was greater than .80.

Tutor evaluations commonly provide both formative and summative feedback. Formative feedback provides tutors with an opportunity to modify their actions during the course to better meet the students' needs. The current study used Hay and Schmuck's (1993) tool to evaluate and provide formative feedback to tutors involved in a PBL occupational therapy program at McMaster University.
The first PBL course that occupational therapy students encounter at McMaster University runs for 14 weeks and focuses on the development of basic knowledge in the biological and behavioral sciences and in the occupational therapy profession. Another objective of this course is to learn "how to learn" in a self-directed, problem-based fashion. The 30 students enrolled in the course participated in the study. All had previously completed a 4-year undergraduate degree but had no experience with a PBL curriculum. All students were women, and their mean age was 25.9 years (SD = 2.4). The participants were randomly assigned to five tutorial groups of six. Each of the five tutors was an experienced tutor; had completed a tutor training course at McMaster University's Faculty of Health Sciences; and was familiar with the problems, procedures, and student evaluations used in the course. All tutors had graduate training, and all but one were experienced occupational therapists.

Tutorial groups met twice weekly in 2.5-hour sessions for 14 weeks. Tutors met weekly with the tutorial coordinator to discuss administrative issues and individual concerns. For the course, each group selected 6 problems to investigate from a bank of 10 possibilities. Problem selection varied according to the learning needs and interests of the individual group. The only constraint was that the group address each of the 25 course objectives at some point.

The students completed the tutor evaluations anonymously at Weeks 6 and 13 and submitted them to the tutorial coordinator. For the Week 6 evaluations, the coordinator tabulated the results for each group and provided summaries to both the tutors and their students before the Week 7 tutorial group session. In addition, each tutor received a summary of results for all tutors combined as well as their own raw evaluation forms. Students received only the summarized results of their tutorial group. The summarized evaluations were discussed openly within each tutorial group. Individual tutor results remained confidential.

The Week 6 evaluations were used for formative purposes. The summarized results for all tutors were discussed during the Week 7 tutors' meeting. These results indicated four specific items that required attention by all tutors: (a) provide insight into the objectives of the course; (b) assist the group in determining when an appropriate knowledge level had been attained; (c) provide consistent, reasonable, verbal feedback; and (d) develop students' individual communication skills. Strategies for remediation were developed. For example, for the first item, the tutors decided to return to the course objectives in subsequent tutorial group sessions and clarify the differences between explicit knowledge objectives (e.g., biological, psychological, sociological) and implicit learning-how-to-learn objectives (e.g., self-directedness, group skills, critical thinking). Tutors were also advised to address areas needing attention in their individual evaluations after comparing them with the summarized evaluation of the group. The areas of concern differed among tutors. For example, two extremely busy tutors were rated poorly regarding enthusiasm about educational role, and they attempted to address this concern within their tutorial groups.

To determine whether these strategies improved tutor performance, a repeated measures analysis of variance was performed on the 30 evaluations completed at both Week 6 and Week 13. The mean scores of each item increased from Week 6 to Week 13, ranging on a 7-point scale from 4.3 to 5.6 (SD = .83-1.3) in Week 6 and from 5.2 to 6.2 (SD = .45-1.5) in Week 13. Eight items showed significant increases (p < .05) in ratings, including the previously noted four items earmarked for attention and (a) show enthusiasm regarding educational role, (b) provide reasonable student autonomy in establishing learning objectives, (c) provide reasonable student autonomy in selecting learning strategies, and (d) develop student problem-solving skills. The two tutors with the lowest mean scores on the enjoyment of educational role item at midterm evaluation had a mean increase of one full scale point in the final evaluation.

Discussion

The tutor evaluation scale proved useful in improving tutor performance within this single tutorial course. None of the four items identified as needing attention at the midterm evaluations had previously been noted as concerns among the tutors. This insight allowed the tutors to respond to student concerns, initiating change during the course rather than recognizing problem areas after its completion. All four of these items had significant increases in the second evaluation.

The three additional items directly related to self-directed, PBL (i.e., establishing learning objectives, selecting learning strategies, developing problem-solving skills) also showed significant increases. An objective of tutors was to encourage the development of skill and comfort in these areas because this was the students' first exposure to this style of learning. The item related to enthusiasm regarding educational role also improved significantly, although it had not been singled out for attention in the tutors' meeting. In interviews after the course, the two tutors...
who had targeted this area stated that they had not realized the effect that their clinical demands had on their groups’ perceptions of their commitment and had made concerted efforts to address this particular item after the midterm evaluations.

Although the evaluation scale we used has proved useful in providing quick insights and allowing the provision of immediate responses, its usefulness in directing long-term changes in tutor performance requires study. The tutors in this study found the scale to be helpful in providing both insight into their performance and a springboard for discussion of their role with the tutorial groups. The evaluation scale and the procedure used allowed students to identify tutor weaknesses and provided tutors the opportunity to respond constructively.

Acknowledgments
I thank the Programme for Educational Development and the members of the Tutor Evaluation Working Group in the Faculty of Health Sciences, McMaster University, Hamilton, Ontario, for their assistance.

References


