

Peer versus staff tutoring in problem-based learning*

HENK SCHMIDT, ARIE VAN DER AREND, IRMA KOKX & LOUIS BOON
*Department of Educational Development and Research, P.O. Box 616, 6200 MD Maastricht,
The Netherlands (*author for correspondence)*

Abstract. Effects of student versus staff tutoring on student learning in a problem-based, health sciences curriculum were studied. Academic achievement of 334 tutorial groups guided by staff tutors was compared with achievement of 400 groups guided by student tutors. In addition, students rated their tutor's performance on four behaviors considered critical to facilitating student learning. Overall, students guided by a staff tutor achieved somewhat better. In terms of practical significance, the difference was, however, fairly small. Staff tutors were rated as more knowledgeable and their contributions as more relevant. In addition, they asked stimulating questions to a larger extent. However, an interaction effect was found between the ratings and the year of study: Peer tutors displayed the supportive behaviors more extensively in the first year, whereas staff tutors' ratings were higher as the curriculum advanced. These results were interpreted in terms of the cognitive congruence framework.

The purpose of this article is to report results of a field experiment in which effects of peer tutoring versus staff tutoring on student learning were studied in the context of a problem-based curriculum. Problem-based learning is an approach to higher education which can be characterized as follows: A collection of carefully constructed problems is presented to small groups of students. These problems usually consist of a description of a set of observable phenomena or events which are in need of explanation. The task of the student group is to discuss these problems and produce tentative explanations for the phenomena described, in terms of some underlying process, principle or mechanism. An example of a problem is the following test: 'A 55-year old woman lies on the floor crawling in pain. The pain emerges in waves and extends from the right lumbar region to the right side groin and the front of the right leg.' Given this text, medical students are required to find explanations for the source of the pain described, its physiological processes and the mechanisms through which it extends itself from a certain area to other areas of the body. It is essential to the method that students' prior knowledge of the problem is, in itself, insufficient to understand it in depth. During initial discussion, dilemmas will arise and questions will come up that can be used

* Parts of this article have been presented to the Annual Meeting of the American Educational Research Association, Atlanta, GA, April, 1993.

as learning goals for subsequent, individual, self-directed learning. While discussing a problem, the group is guided by a tutor, usually a member of the faculty. His or her task is to stimulate the discussion, to provide students – if necessary – with some subject-matter information, to evaluate progress being made and to monitor the extent to which each group member contributes to the group's goals (Barrows & Tamblyn, 1980; Schmidt, 1983; Norman & Schmidt, 1992).

From this brief description, it may be clear that the tutor plays an important role in encouraging and guiding students in their learning. This is not an easy task, because the range of the tutor's possible actions is to a large extent determined by the situation. What the tutor may contribute depends on the kind of difficulties students encounter while working on the problem-at-hand. Teachers cannot fully prepare themselves for this kind of teaching and, therefore, do not always feel comfortable with this role. By contrast, the role with which most teachers feel familiar – lecturing – is not in high demand in tutorial groups because, in problem-based learning, students are required to collect the necessary information largely by themselves. Consequently, academic staff are often reluctant to act as tutors. This state of affairs constitutes a reason to look for alternatives to academic staff as tutors. It was expected that students, who themselves had experienced problem-based learning, would have less trouble adapting to the tutor role and would be more sensitive to the specific needs of tutorial groups in terms of support, information and guidance.

A second reason for involving older students as tutors in problem-based curricula is of a didactic nature. There is a literature (e.g. Collier, 1980; Cornwall, 1980) suggesting that, in some respects, peers would be more suited to tutoring other students because there would be more congruence between their cognitive structures and those of the students tutored. This cognitive congruence would qualify the student tutors to better understand the difficulties encountered by the tutees and, hence, would enable them to respond in a more adequate fashion. On the other hand, however, student tutors clearly have less subject-matter knowledge than staff tutors and would, from this perspective, be less able to help other students, because helping would imply a 'pooling of ignorance' as De Volder, De Grave & Gijssels (1985) put it.

The existing literature on the effects of peer versus staff tutoring in higher education is scattered and its results are generally inconclusive. Carsrud (1979), for instance, found that students guided by a staff teacher achieved better than student-tutored groups. Bloxom, Caul, Fristoe & Thomson (1975) and De Volder et al. (1985), however, found no differences. In a recent review, Moust (1993) suggested that one of the reasons for inconsistent findings in

peer tutoring studies may be lack of statistical power, since many of these studies were carried out with only small number of groups.

In order to study the effects of student versus staff tutors on student learning, a large-scale experiment was conducted. Academic achievement of 340 tutorial groups guided by staff tutors was compared with achievement of 400 groups guided by student tutors. In addition, students rated their tutor's performance on four behaviors considered critical to facilitating student learning.

Method

Subjects

Subjects were exactly 1800 students of health sciences at the University of Limburg, the Netherlands. In addition, 411 tutors were involved in the program; 152 staff tutors and 259 student tutors. Most of these tutors ran more than one tutorial group. Students were on average a member of five different tutorial groups. For each six-week course, both students and tutors were assigned to a different tutorial group in an entirely random fashion.

Description of the health sciences curriculum

The four-year health science curriculum consists of a large number of courses of equal length. Data from 98 courses taught in the academic year 1989–1990 were included in the analysis, which is 82% of the total number of courses. For 21 courses insufficient data were available. The courses were taught following the same general problem-based format: students met with their tutor in small-group tutorials twice a week for two hours. An average of six hours of additional activities were scheduled, such as labs, skills training and occasional lectures. The remainder was spent on self-study.

Procedure

At the end of each course, a questionnaire was administered to all students. It inquired about various aspects of the course, including the tutor's behavior. Four items are of relevance to the purpose of the present study: 'The tutor made use of his or her subject-matter knowledge to help us', 'The tutor's questions stimulated the discussion', 'The contributions of the tutor were relevant' and 'At regular intervals, the tutor evaluated with us the group's functioning'. The student could indicate her or his impression on a five-point Likert scale ranging from 'entirely disagree' to 'entirely agree.' An

achievement test was administered after each course and students received a mark on a ten-point scale ranging from 1 to 10; 6 being the pass score. The test either consisted of short-essay questions or multiple-choice items. Tutors were not involved in assessing their students' knowledge level.

Analysis

Data were aggregated at the level of tutorial groups. This was done because one cannot assume that scores produced by members of an interacting group are completely independent of each other (Cohen, 1980; Howard & Maxwell, 1980). Although each student was tested several times during the academic year and contributed to the aggregated scores of several tutorial groups, data at the tutorial group level can nevertheless be considered independent, since both students and tutors were randomly assigned to these groups.

The data were analyzed using ANOVA with staff versus peer tutor and year of study as independent variables. The latter variable was included in the analysis because informal observation suggested that effects of tutors on the students' learning to some extent may be dependent on amount of the students' experience with the problem-based format (here expressed as number of years in the curriculum).

Results

Achievement

In most curriculum years, students working under the guidance of staff tutors received somewhat higher marks. The effect of staff versus peer tutoring was small – average score in the staff tutoring condition: 6.79; peer tutoring: 6.64. This difference, however, is statistically significant, $F(1,733) = 9.42$, $p < 0.002$. The effect of curriculum year was also significant, $F(3,731) = 3.26$, $p < 0.02$. The interaction between both variables was however non-significant: $F(3,731) = 0.60$, $p < 0.62$. The error term was equal to 0.44. These results indicate that peer and staff tutors do indeed have a (small) differential impact on the achievement of students. Average scores as a function of tutor background and curriculum year are given in Table 1. Pairwise contrasts analyses revealed that differences in year 1 and 4 are nonsignificant.

Table 1. Average achievement of health sciences students as a function of tutor background and curriculum year (Number of groups on which the average is based between brackets).

Curriculum year	1	2	3	4	Total
Peer tutor	6.56 (115)	6.65 (153)	6.69 (117)	6.66 (15)	6.64 (400)
Staff tutor	6.62 (84)	6.86 (133)	6.88 (84)	6.69 (33)	6.79 (334)

Tutor behavior

Students rated their tutors on four critical behaviors: To what extent did the tutor use his subject-matter knowledge to help the group? Were his contributions relevant? Did he ask questions that stimulated discussion and did he evaluate the group's progress?

Staff tutors made a more extensive use of their subject-matter knowledge than peer tutors: $F(1,744) = 24.52$, $MSe = 0.56$, $p < 0.001$. The influence of year of study was negligible. However, a significant interaction effect was found: $F(3,742) = 8.02$, $p < 0.001$, indicating that staff tutors used their knowledge not to the same extent in each curriculum year. Inspection of the average scores reveals that, in fact, peer tutors are rated higher in the first year, whereas staff tutors get higher ratings in the three subsequent years.

The latter phenomenon was also observed in the data concerning the relevance of the tutor's contributions. Both a main effect of staff versus peer tutoring, $F(1,744) = 3.94$, $MSe = 0.34$, $p < 0.05$, and a significant interaction effect were found: $F(3,742) = 6.97$, $p < 0.001$. Again, peer tutors were rated higher in the first year, whereas staff were rated higher in subsequent years.

An interaction effect was also found for 'Asking stimulating questions,' $F(3,742) = 3.64$, $MSe = 0.44$, $p < 0.05$. In the first year, peer tutors are doing better in this respect whereas in subsequent years, staff tutors are doing a better job. There was no main effect for this variable; overall, peers and staff displayed this behavior to the same extent.

A different pattern, however, emerged with respect to 'Evaluating the group's functioning.' Peer tutors evaluated more extensively than staff tutors: $F(1,744) = 38.68$, $MSe = 0.56$, $p < 0.001$. However, in the course of the years, this behavior becomes less prominent: $F(3,744) = 4.97$, $p < 0.01$, both for peer and staff tutors, since no interaction effect was found.

Of course, these ratings may not be entirely independent. A common factor may underlie the behaviors assessed. However, confirmatory factor analysis

showed that neither a one-factor, nor a two factor, hypothesis fitted the data entirely satisfactorily. The one-factor hypothesis produced a Chi-square of 91.018, $df = 2$, $p < .0001$. Bentler-Bonett's Comparative Fit Index for this case equaled 0.94. The two factor solution assumed that asking stimulating questions and evaluating the group's functioning represented a "process" dimension, while making use of one's subject-matter knowledge and the relevance of the tutor's contributions could be considered a "content" factor. For this case Chi-square equaled 42.996, $df = 1$, $p < 0.001$; and Bentler-Bonett's Comparative Fit Index was equal to 0.97. These findings seem to indicate that the tutor ratings do not entirely tap the same underlying trait.

Discussion

Tutors are to facilitate their student's learning in problem-based curricula in an indirect way. Since they are not supposed to lecture, they must help students through a number of supportive behaviors, such as asking stimulating questions and evaluating the group progress at regular intervals (Peterson & Swing, 1985). It is assumed that these behaviors are particularly effective when tutors have sufficient subject-matter knowledge relative to the issues that students deal with in the tutorial group and actually use that knowledge while guiding students (Moust, 1993). Finally, it is expected that these behaviors would eventually lead to superior performance by their students relative to students guided by less effective tutors. Since peers may be considered less knowledgeable than academic staff, students supported by staff may be expected to achieve better. The present large-scale field experiment provided support for these assertions. Staff tutors were judged to display the required behaviors to a greater extent than student tutors. They contributed more, using their subject-matter knowledge, their contributions were considered more relevant, and they asked stimulating questions to a greater extent than did student tutors. (They did, however, evaluate the group's progress less frequently than students.) In addition, a small but statistically significant effect on achievement was found, carrying the suggestion that these behaviors indeed were effective.

Interestingly, the first curriculum year constituted an exception to the overall trend. Here student tutors were judged as generally contributing more to the students' learning. In addition, and contrary to the overall findings, no difference in achievement was found in this curriculum year between groups under both conditions of the experiment. Moust (1993) suggests that, in the first year, students prefer tutors that behave 'cognitively congruent,' that is: display an understanding of the way in which students think and express them-

selves in their language. The subject-matter expertise of the tutor becomes more important only over the course of the years.

References

- Barrows, H.S. & Tamblyn, R.M. (1980). *Problem-based Learning*. New York, NY: Springer Publishing.
- Bloxom, B., Caul, W.F., Fristoe, M. & Thomson, W.J. (1975). On the use of student led discussion groups. *The Educational Forum* 39: 223–230.
- Carsrud, A.L. (1979). Undergraduate tutors: Are they useful? *Teaching of Psychology* 6: 46–49.
- Cohen, P.A. (1980). Student ratings of instruction and student achievement: A meta-analysis of multisection validity studies. *Review of Educational Research* 51: 281–309.
- Collier, G.K. (1980). Peer group learning in higher education: The development of higher order skills. *Studies in Higher Education* 5: 55–62.
- Cornwall, M.G. (1980). *Students as Teachers: Peer Teaching in Higher Education*. Amsterdam, The Netherlands: Centrum voor Onderzoek van het Wetenschappelijk Onderwijs.
- De Volder, M.L., De Grave, W.S. & Gijsselaers, W. (1985). Peer teaching: Academic achievement of teacher-led versus student-led discussion groups. *Higher Education* 14: 643–650.
- Howard, G.S. & Maxwell, S.E. (1980). Correlation between student satisfaction and grades: A case of mistaken causation? *Journal of Educational Psychology* 72: 810–820.
- Moust, J.H.C. (1993). *De rol van tutoren in probleemgestuurd onderwijs: contrasten tussen student- en docent-tutoren* (The role of tutors in problem-based learning: Contrasts between student and staff tutors) Maastricht, The Netherlands: Universitaire Pers Maastricht.
- Norman, G.R. & Schmidt, H.G. (1992). The psychological basis of problem-based learning: A review of the evidence. *Academic Medicine* 67: 557–565.
- Peterson, P.L. & Swing, S.R. (1985). Students' cognitions as mediators of the effectiveness of small-group learning. *Journal of Educational Psychology* 77: 299–312.
- Schmidt, H.G. (1983). Problem-based learning: Rationale and description. *Medical Education* 17: 11–16.