

## Is Tutor Performance Dependent on the Tutorial Group's Productivity?: Toward Further Resolving of Inconsistencies in Tutor Performance

Diana H. J. M. Dolmans, Ineke H. A. P. Wolfhagen,  
Ron J. I. Hoogenboom, and Cees P. M. van der Vleuten

Department of Educational Development and Research  
Maastricht University  
Maastricht, The Netherlands

**Background:** Many studies have been conducted on tutor performance in problem-based curricula. In the past, the implicit assumption behind these studies was that tutor performance is a relatively stable characteristic. More recent studies demonstrate that a tutor's performance may be dependent on other circumstances, such as the level of structure in the curricular materials. The aim of this study was to investigate whether a tutor's performance is also dependent on the tutorial group's productivity.

**Purpose:** The idea is that low-productive tutorial groups require much more input from a tutor than high-productive groups. In the problem-based curriculum under investigation, most tutors guide 2 tutorial groups within the same unit. A salient finding in this problem-based curriculum was that some tutors who guide 2 tutorial groups within the same unit have discrepancies in their tutor performance across the 2 groups. This finding might be explained by differences in both tutorial groups. In this study, first the scope of the discrepancy phenomena was studied. Second, the relation between the tutor's performance and the tutorial group's productivity was studied.

**Methods:** The data set for this study included 136 tutors who, in total, ran 272 tutorial groups (each tutor ran 2 groups per unit). The analyses were conducted at the tutorial group level. Students were asked to judge the performance of their tutor. Low, medium, and high levels of tutor performance were distinguished. Tutors who were qualified as "low level of performance" in one tutorial group and "medium level of performance" in the other tutorial group were considered to have a discrepancy in their tutor performance: "discrepancy tutors." The same holds for tutors with medium level of performance in one group and high level of performance in the other group or low level of performance in one group and high level of performance in the other group. All other tutors were considered "nondiscrepancy tutors." The nondiscrepancy tutors had equal levels of performance in both groups: a low, medium, or high level. For each type of tutor (discrepancy tutors and nondiscrepancy tutors) the average tutorial group's productivity score was computed.

**Results:** The results show that 39% of the tutors were classified as discrepancy tutors. In addition, it was found that a discrepancy tutor with a low level of tutor performance in one group also had a low productivity score in this group, whereas a high level of tutor performance corresponds with a high level of the tutorial group's productivity. Furthermore, the results show that nondiscrepancy tutors with a high level of tutor performance receive high tutor performance scores, irrespective of the tutorial group's level of productivity.

**Conclusions:** These findings demonstrate that the tutorial group's productivity is another influencing factor in determining tutor performance. Low-productive groups require much more input from a tutor than high-productive groups. Nondiscrepancy tutors with consistent low levels of tutor performance and discrepancy tutors lack certain competencies that are needed when being confronted with a low-productive

*tutorial group. Nondiscrepancy tutors with a high level of tutor performance, on the contrary, know how to deal with low-productive tutorial groups, due to which their tutor performance is high irrespective of the tutorial group's productivity. Thus, a tutor's performance seems to be part tutor specific and part situation specific (i.e., dependent on the group's productivity).*

Many studies have been conducted on tutor performance in problem-based curricula.<sup>1-8</sup> These studies usually focus on the influence of a tutor's expertise on student performance. The implicit assumption behind these studies is that tutor performance is a relatively stable characteristic and is highest when the subject matter of the unit in which the tutor is involved is related to his or her expertise. More recent studies reveal, however, that a tutor's expertise only has an impact on student performance if a unit does not match the students' prior knowledge or when the unit is less structured.<sup>5,9</sup> This indicates that tutor performance might not be a stable characteristic but may be dependent on other circumstances as well. The circumstances investigated so far deal with students' level of prior knowledge and the level of structure in the unit. Another influencing factor might be the group's effectiveness. A low-productive group requires much more input from a tutor than a high-productive group.

A few years ago, the educational committee of the Medical School of the Maastricht University decided that teachers involved as tutors in a problem-based unit should also be involved in the construction of that particular unit. This policy required that each teacher involved in the construction of a particular unit guide two tutorial groups during the flow of the unit. The idea was that a tight match of both activities would improve the teacher's engagement with the unit and as a consequence would result in better tutor performance. A concomitant circumstance was that the performance of a tutor within the same unit could be compared for two different tutorial groups. A salient finding that emerged immediately after implementation of this policy was the discrepancy in tutor judgments by students at the end of the unit. One tutorial group, for instance, judged a tutor's level of performance as high, whereas the second tutorial group judged the tutor's level of performance within the same unit as relatively low or medium. This finding undermines the idea that the performance of a tutor is a relatively stable characteristic.

Explanations given by tutors for these remarkable discrepancies were mostly related to the tutorial group's low level of productivity, caused by lack of student effort and problems of group dynamics. Most tutors assumed the perceived low level of their tutor performance to be a halo-effect; that is, if a tutorial group is perceived by students as low productive, the tutor is assumed to be performing at a low level and will as a consequence receive a relatively low qualification. An

alternative explanation is that a low-productive tutorial group requires much more input from a tutor or requires interventions that some tutors do not possess. Some tutors do not have the skills to intervene adequately, resulting in a discrepancy in their level of tutor performance score across two groups, whereas others possess these skills, yielding high levels of tutor performance evaluations in both tutorial groups, irrespective of the group's productivity. Hitchcock and Anderson,<sup>10</sup> concluded that some tutors have found themselves in dysfunctional groups without the skills necessary to foster positive group development. These researchers conducted a study to generate strategies tutors can use to deal with dysfunctional groups. It is assumed in this study that tutors with discrepancies in their tutor performance scores probably do not know how to intervene to foster positive group productivity.

In a study by Gijsselaers and Schmidt,<sup>11</sup> investigating the relation between tutor performance (independent variable) and the tutorial group's performance (dependent variable), it was shown that 32% of the variance in group functioning could be explained by the variance in tutor performance. In this modeling approach to learning in problem-based curricula, tutor performance is assumed to be an independent variable influencing the tutorial group's productivity.<sup>11-13</sup> It is, however, self-evident that in reality tutor performance and tutorial group performance influence each other mutually and that tutor performance may partly be dependent on the tutorial group's productivity. The phenomenon of a discrepancy in level of tutor performance for two tutorial groups within the same unit was used as a starting point for investigating this relation. In the situation under study, the tutor's expertise did not differ in relation to the unit content, nor did the structure of the curriculum alter the relation, because each tutor ran two tutorial groups within one unit. This setting provided an excellent opportunity to investigate inconsistencies in tutor performance.

The first aim of this study was to investigate the scope of the discrepancy phenomena. Second, we sought to investigate the relation between the tutor's level of performance and the tutorial group's level of productivity for tutors with and without a discrepancy in tutor judgments by students across two tutorial groups within the same unit. In other words, the aim was to assess whether tutor performance may be partially situation specific (i.e., dependent on the group's productivity).

**Table 1.** The Number of Tutors With or Without a Discrepancy Score in Their Tutor's Performance, Their Mean Tutor Performance Scores, Standard Deviation, and Number of Tutors (N) and Their Corresponding Tutorial Group's Productivity Score

	No. of Tutors <sup>a</sup>		Tutor Performance Score <sup>b</sup>			Productivity Score <sup>c</sup>		
	N	%	M	SD	N	M	SD	N
<b>Nondiscrepancy Tutors</b>								
Two Low Scores (Low Level)	29	21	6.67	0.69	58	3.00	0.67	58
Two Medium Scores (Medium Level)	18	13	7.67	0.13	36	3.40	0.58	36
Two High Scores (High Level)	36	27	8.40	0.34	72	3.88	0.51	72
Total	83	61						
<b>Discrepancy Tutors</b>								
One Low and One Medium Score	18	13	7.23	0.58	36	3.34	0.58	36
One Low and One High Score	15	11	7.49	0.72	30	3.41	0.74	30
One Medium and One High Score	20	15	7.99	0.35	40	3.68	0.43	40
Total	53	39						

<sup>a</sup>N = 136. <sup>b</sup>1-10. <sup>c</sup>1-5.

**Table 2.** Mean Score for the Tutor's Performance, the Corresponding Standard Deviation, Number of Tutors (N), and the Tutorial Group's Productivity Score for Tutors with a Discrepancy in Their Tutor Performance Score for Each Tutorial Group Separately

	Tutor Performance Score <sup>a</sup>			Productivity Score <sup>b</sup>		
	M	SD	N	M	SD	N
<b>One Low and One Medium Tutor Score</b>						
Low Score	6.79	0.50	18	3.12	0.61	18
Medium Score	7.68	0.10	18	3.56 <sup>c</sup>	0.48	18
<b>One Low and One High Tutor Score</b>						
Low Score	6.87	0.46	15	2.87	0.43	15
High Score	8.11	0.19	15	3.95 <sup>d</sup>	0.59	15
<b>One Medium and One High Tutor Score</b>						
Medium Score	7.72	0.16	20	3.54	0.41	20
High Score	8.27	0.26	20	3.81 <sup>e</sup>	0.42	20

<sup>a</sup>1-10. <sup>b</sup>1-5. <sup>c</sup>The average productivity scores (3.12 and 3.56) differ significantly,  $F(1, 35) = 5.86, p = .02$ . <sup>d</sup>The average productivity scores (2.87 and 3.95) differ significantly,  $F(1, 29) = 33.01, p = .00$ . <sup>e</sup>The average productivity scores (3.54 and 3.81) differ significantly,  $F(1, 39) = 4.44, p = .04$ .

scores. The results are presented in Table 2. As can be seen, for tutors with a discrepancy in their performance, a high level of tutor performance corresponds with a relatively high productivity score and a low level of tutor performance corresponds with a relatively low productivity score. The average productivity scores of both tutorial group's differed for all types of discrepancy tutors significantly ( $p < .05$ ), as can also be seen in Table 2.

To further investigate the relation between tutor performance and the tutorial group's productivity, the productivity score per tutorial group was classified as relatively low, medium, or high. For the purpose of our analyses, these classifications were assigned the following values: low = 1, medium = 2, and high = 3. The hypothesis tested was whether a tutor characterized as a low level of tutor performance, also had a low productivity score, and a tutor characterized as a high level of tutor performance, also had a high productivity score. In other words, is a tutor's performance score relatively consistent with the group's productivity score? If this hypothesis could be confirmed, it would imply that the

tutor's performance score minus the group's productivity score would result in an average point of difference that is equal to zero. The results are presented in Table 3. The average point of difference for all nine types of tutors included in Table 3 differs significantly,  $F(8, 271) = 12.18, p = .00$ . As can be seen, the average point of difference is equal to zero only for the nondiscrepancy tutors with a medium level of tutor performance. The nondiscrepancy tutors with a low level of tutor performance and the discrepancy tutors with a low level of tutor performance have a negative point of difference, which implies that, on average, the group's productivity score is higher than the tutor's performance score. The nondiscrepancy tutors with a high level of tutor performance and the discrepancy tutors with a high level of tutor performance have a positive point of difference, which implies that, on average, the group's productivity score is lower than the tutor's performance score. Thus, a nondiscrepancy tutor with a high level of tutor performance is a tutor with high tutor performance scores, irrespective of the circumstance of the tutorial group's productivity. A nondiscrepancy tutor with a low level of

prior knowledge in relation to the unit content and the level of structure in the curricular materials. This study demonstrates that the tutorial group's productivity is another influencing factor, due to which a tutor's performance is not a stable characteristic. Tutorial groups with relatively low levels of productivity require much more input from a tutor than high-productive groups, due to which a tutor's competencies are more likely to be actually needed by students. Nondiscrepancy tutors with consistent low levels of tutor performance and discrepancy tutors lack certain competencies that are needed when dealing with low productivity groups.

The conclusion that a tutor's performance is not a stable characteristic, but is rather situation specific, should, however, be taken with caution for three reasons. First, the results of this study show that fluctuation in tutor performance scores was not found for all tutors, but for 39% of the tutors involved in this study. Second, it was found that tutors with high levels of tutor performance score more or less stably; they receive high tutor performance scores irrespective of the tutorial group's productivity. Third, in a previous study it was shown that tutor performance over repeated occasions of evaluation was rather stable.<sup>8</sup> Thus, a tutor's performance seems to be part tutor specific and part situation specific and is not more situation specific than tutor specific.

Finally, further investigations should be conducted to discover what kinds of skills a tutor should possess to be able to deal with tutorial groups with relatively low levels of productivity. It is also important to determine better what low productivity in a learning group means. An attempt in this area is a recent study in which it was demonstrated that the group's productivity is closely related to motivational dimensions and cognitive dimensions, (i.e., cohesion, quality of interaction, and elaboration).<sup>15</sup> It would then be useful to determine whether using certain skills, thus identified, would yield less discrepancy in tutor performance and higher levels of tutor performance. Discrepancy tutors could, for example, be offered the opportunity to participate in training sessions like the one developed by Hitchcock and Anderson.<sup>10</sup> These sessions were aimed at generating strategies tutors can use to deal with dysfunctional groups. Subsequently one could investigate whether such training would yield less discrepancy in their tutor performance scores.

## References

1. Davis WK, Nairn R, Paine ME, Anderson RM, Oh MS. Effects of expert and non-expert facilitators on the small-group process and on student performance. *Academic Medicine* 1992;67:470-74.
2. Davis WK, Oh MS, Anderson RM, Gruppen L, Nairn R. Influence of a highly focused case on the effect of small-group facilitators' content expertise on students' learning and satisfaction. *Academic Medicine* 1994;69:663-69.
3. Eagle ChJ, Harasym PH, Mandin H. Effects of tutors with case expertise on problem-based learning issues. *Academic Medicine* 1992;67:465-69.
4. Schmidt HG, Arend A van der, Moust JHC, Kokx I, Boon L: Influence of tutor's subject-matter expertise on student effort and achievement in problem-based learning. *Academic Medicine* 1993;68:784-91.
5. Schmidt HG. Resolving inconsistencies in tutor expertise research: Does lack of structure cause students to seek tutor guidance? *Academic Medicine* 1994;69:656-62.
6. Silver M, Wilkerson LA. Effects of tutors with subject expertise on the problem-based tutorial process. *Academic Medicine* 1991;66:298-300.
7. Wilkerson LA. The next best thing to an answer about tutors' content expertise in PBL. *Academic Medicine* 1994;69:646-8.
8. Dolmans DHJM, Wolfhagen HAP, van der Vleuten CPM. Long-term stability of tutor performance. *Academic Medicine* 1996;71:1344-7.
9. Dolmans DHJM, Wolfhagen HAP, Schmidt HG. Effects of tutor expertise on student performance in relation to prior knowledge and level of curricular structure. *Academic Medicine* 1996;71:1008-11.
10. Hitchcock MA, Anderson, AS. Dealing with dysfunctional tutorial groups. *Teaching and Learning in Medicine* 1997;9:1:19-24.
11. Gijsselaers WH, Schmidt HG. Development and evaluation of a causal model of problem-based learning. In AM Nooman, HG Schmidt, ES Ezzat (Eds.), *Innovation in medical education: An evaluation of its present status* (pp. 95-113). New York: Springer, 1990.
12. Schmidt HG, Moust JHC. What makes a tutor effective? A structural-equations modeling approach to learning in problem-based curricula. *Academic Medicine* 1995;70:708-14.
13. Schmidt HG, Dolmans D, Gijsselaers WH, Des Marchais, JE. Theory-guided design of a rating scale for course evaluation in problem-based curricula. *Teaching and Learning in Medicine* 1995;7:82-91.
14. Dolmans DHJM, Wolfhagen HAP, Schmidt HG, Van der Vleuten CPM. A rating scale for tutor evaluation in a problem-based curriculum: Validity and reliability. *Medical Education* 1994;28:550-8.
15. Dolmans, DHJM, Wolfhagen, HAP, Van der Vleuten CPM. Motivational and cognitive processes influencing tutorial groups. *Academic Medicine* 1998;73:S22-4.

Received 29 June 1998

Final revision received 23 March 1999