

The effect of directive tutor guidance in problem-based learning of statistics on students' perceptions and achievement

Luc Budé · Tjaart Imbos · Margaretha W. J. v. d. Wiel · Nick J. Broers · Martijn P. F. Berger

Published online: 20 March 2008
© The Author(s) 2008

Abstract In this study directive tutor guidance in problem-based learning (PBL) of statistics is investigated. In a quasi experiment in an educational setting, directive guiding tutors were compared with tutors in a more traditional role. Results showed that the subjective perceptions of the students with regard to the course, the tutor, and the discussions in the tutorial meetings were more positive in the guided condition. The quality of the problems used in the meetings and general tutor functioning were evaluated as equal in both conditions. Achievement was marginally higher in the guided condition. It can be concluded that directive tutor guidance is an effective addition to PBL of statistics.

Keywords Guided learning · Human tutoring · Problem-based learning · Statistics education · Students' achievement

Introduction

In this study the role of the tutor in problem-based learning (PBL) of statistics is investigated. The benefit of human tutoring has been demonstrated in many studies (Chi 1996; Chi et al. 2001; Graesser et al. 1995). Positive aspects of tutor performance in problem-based learning (PBL) have been studied extensively too (Barrows 1988; Dolmans et al. 2002; Dolmans and Wolfhagen 2005; Schmidt and Moust 1995). Summarising these findings it can be concluded that tutoring is effective:

- 1) When tutors and students interact; tutor's actions that prompt for the co-construction of knowledge are positively correlated with deep understanding (Chi 1996, Chi et al. 2001).
- 2) When students are activated; understanding will be improved by answering why, how, and what-if questions (Graesser et al. 1996).

L. Budé (✉) · T. Imbos · M. W. J. v. d. Wiel · N. J. Broers · M. P. F. Berger
Department of Methodology and Statistics, Maastricht University,
P.O. Box 616, 6200 MD Maastricht, The Netherlands
e-mail: Luc.Bude@stat.unimaas.nl

- 3) When self-study is stimulated; self-study will induce higher achievement (Schmidt and Moust 1995, 2000).

One-on one tutoring can be defined as a way of instruction, which is characterised by an interactive and continuous stream of exchanges between a tutor and a student (Chi 1996). Although tutors in PBL operate in a group, their role is more or less the same as in one-on one tutoring. Instead of dispensing knowledge they should try to activate students, stimulate group processes, try to create an atmosphere in which students can optimally participate in the discussions, help students to monitor their own learning, and to stimulate self-study (Schmidt and Moust 2000; De Grave et al. 1999). In PBL both content expert faculty tutors and non-content student tutors are used. Content expert tutors tend to make more subject matter contributions to the discussions than non-content expert tutors. This has been shown to improve especially novice student's performance. It seems that a tutor's expertise can compensate for lack of prior knowledge of students (Schmidt et al. 1993, Dolmans et al. 2002). Research has also shown that tutorial groups with relatively low levels of productivity require more input from a tutor (Dolmans and Wolfhagen 2005). To be able to make subject matter contributions the tutors in PBL usually have tutoring instructions at their disposal. These instructions consist of general information about the course, the main topics, more specific information about the goal of the problems, which subject matter is supposed to be discussed and how to tackle the problems. In the intervention of the current study the tendency of tutors to intervene in the content of the discussion, was taken a step further.

The intervention consisted of additional tutor guidance during the meetings. That is, the tutors more actively guided the discussions in a directive way than it is usual in PBL. The research question of our study was concerned with the effect of this extra directive tutor guidance on students' achievement and their subjective perception of several course aspects, group functioning, the relevancy of tutor contributions, general tutor functioning, and the quality of the problems that were used for the discussions in the tutorial sessions.

In the following we will first give a short description of PBL. Next, we will describe some characteristics of statistical knowledge in relation to the learning of statistics and we will describe our intervention. After that, we will present our study.

Problem-based learning

PBL refers to a variety of approaches to instruction, which all have in common that much of the learning and instruction is anchored in concrete problems (Hmelo and Evensen 2000). A problem can be anything that raises questions germane to the subject matter and affords free inquiry by students (Barrows 1986). For example, a problem can be a patients' case (in medical education), the outcome of a study, a hypothesis about a real life phenomenon, or a statistical problem that students need to solve. A problem that could be used in PBL of statistics is the question of what an appropriate analysis technique would be for a given dataset. While exploring and discussing the problem in an initial tutorial session, students extract key information about the problem and discover deficiencies in their knowledge (Hmelo and Evensen 2000). Based on these knowledge gaps and out of an intrinsic curiosity, students formulate their own learning goals and decide on what they are going to study. Instead of merely being exposed to information, the students in PBL are actively engaged in gathering information that is processed in relation to the presented problem. After the initial tutorial session the students in most PBL courses individually

study relevant literature and in the following session they report back to the group what they have learned.

During the discussions in the tutorial sessions there are two kinds of guidance. Firstly, the problems used in PBL usually provide a fair amount of guidance to the students. Ideally, a problem directs students to the main topics and focuses on important issues. As a consequence the effectiveness of PBL is closely related to the quality of the problems (Albanese and Mitchel 1993; Schmidt and Moust 2000). Secondly, effective tutoring, as explained in the introduction, also involves appropriate guidance from the tutors in the discussions of the subject matter. This guidance usually consists of hinting the students, pointing to relevant topics, and helping students to monitor their own learning. Tutors traditionally ask questions like: can you explain this; do you understand that; can you see why that is important, etc? However, if a tutor contributes too much to the discussion in PBL, self-study time decreases (Schmidt and Moust 2000). Moreover, research has shown that education in general should not be too directive. Otherwise, students may lose the idea of having control, which may cause autonomy and self-regulated learning to decrease, which on its turn may cause students to get less motivated and become more passive (Ames 1992; Deci and Ryan 1985; Lepper et al. 1997; Pintrich 2003). Passiveness is unfavourable, because the best way to acquire knowledge is to actively process the subject matter (see e.g. Anderson 1983; Graesser et al. 1996; Chi et al. 1989).

In spite of these findings we instructed the tutors in the intervention condition of our study to guide the discussions in a directive way. We think that in PBL of statistics such an approach might have beneficial effects on students' understanding of the subject matter, for two reasons. First, students have little prior knowledge of the subject. Second, statistics is a special discipline and requires special teaching, as will be explained in the next paragraph.

Characteristics of statistical knowledge

University teaching shows differences among disciplines, which may be due to genuine differences among fields of learning, such as the natural sciences and social sciences (Smeby 1996). Students in natural sciences seem to receive more guidance and structured supervision than students in the social sciences (Acker et al. 1994).

Statistics is a field of learning in which the concepts are relatively abstract and interconnected to a great degree, which makes it difficult for students to understand (Gal and Garfield 1997; Schau and Mattern 1997). For example, the fact that the mean is sensitive to outliers can affect the interpretation of the results of an analysis of variance. This example also illustrates that statistical knowledge is hierarchical. A student first has to know what the mean is, then has to understand the concept deviation from the mean, before variance can be understood, which has to be comprehended before analysis of variance can be understood. Statistical knowledge is said to be highly structured, because of this interconnectivity of the concepts and the hierarchy of the knowledge.

In abstract and highly structured knowledge domains (statistics, computer programming, and mathematics) research has shown that learning can be improved when teaching is more directive. For example, by using worked examples in mathematics (Carroll 1994; Sweller 1999; Sweller and Cooper 1985), and computer programming (Tuovinen and Sweller 1999); guided discovery in computer programming (Debowski et al. 2001; Fay and Mayer 1994; Lee and Thompson 1997); and by explicating strategies in statistics (Paas 1992). The interventions in all these studies acted as a guide, showing the students the most

effective way to achieve their aim. In all studies it was shown that the performance of students improved. Our intervention intended to have the same effect.

Directive tutor guidance

As explained in the paragraph on PBL, students in PBL are directed by the problems and the tutors. It was expected that for statistics education extra directive guidance provided by the tutor would promote students' understanding and that students would gain a better insight into the content of the course. In the current study the tutors in our intervention condition provided directive guidance by asking questions. They received a detailed list of specific questions, in addition to the more general tutoring instructions with background information that tutors usually have. By asking questions, the tutors in the intervention condition initiated the discussion and they directed the discussion in a predetermined way. The tutors intervened as soon as the students would go astray or omit an important subject. The questions were based on how experts would deal with the presented problems. Step by step, in a subscribed order the tutoring questions guided the students through the subject matter. The guiding questions focused the discussions on relevant issues, indicating the direction of reasoning without hampering the active learning processes. Tutors received written directions and training in how to use the instructions. They were instructed to specifically ask for the relations between the concepts. This should make more explicit the interconnectivity of the concepts. The tutors were also instructed to make sure that the topics were discussed in a prescribed order. This was done to adhere to the hierarchy of the knowledge. We will refer to this intervention as directive tutor guidance. In this study a condition in which tutors provided this directive guidance was compared to a standard (not guided) condition. In the standard condition the tutors were involved in facilitating the group and the discussion, but not directing it.

It was hypothesised that the directive tutor guidance would have a positive effect on students' achievement and subjective perception of the course, group functioning, and the relevancy of tutor contributions. With regard to students' subjective perception of general tutor functioning, a negative effect was expected due to the reduced autonomy of the students. No effect was anticipated with regard to the quality of the problems that were used for the discussions in the tutorial sessions.

Method

Participants

Two hundred and six students, enrolled in a bachelor statistics course of Health Sciences at the University of Maastricht, participated in this study. They were randomly assigned to 24 tutorial groups. These 24 tutorial groups were randomly assigned to 14 tutors (ten tutors had two groups, four had one group). Finally, the tutors were assigned to either the guided condition or to the control condition using blocked randomisation. This resulted in 12 groups ($N = 102$) being assigned to the guided condition and 12 groups ($N = 104$) to the control condition. The tutors were members from the department of Epidemiology and the department of Methodology and Statistics. All tutors had sufficient knowledge of the subject matter and had minimal 3 years experience in tutoring.

Materials and procedure

The bachelor statistics course in which this study was conducted took seven weeks. Students discussed the topics in a weekly tutorial group meeting on the basis of a list of problems. The main topics were methodological and statistical subjects. The methodological topics were randomised clinical trials and quasi-experimental designs. The statistical topics were the central limit theorem, *t*-tests, ANOVA, and linear regression analysis. For an example of a problem see Appendix A. In the control condition, tutors followed the usual PBL tutoring procedure, facilitating the group and learning processes, but without directing the discussions. In the intervention condition, the tutors used a list of questions to guide the discussion in the tutorial group meetings in a directive way. The list consisted of questions for each statistical problem that was discussed in the tutorial meetings. See Appendix B for an example. The questions were in a prescribed order, oriented toward the main topics and their relations. Typically, in PBL the discussion starts with the problem definition. The tutor would therefore start with the question what the problem is about and successively guide the students through the subject matter. The questions were used in the initial discussion to make sure that all relevant topics were covered in the correct order and students would see what they had to know about the subject matter and what they did not yet know. This enabled the formulation of correct learning goals. After individual study, in the reporting phase, the questions were used during the discussions to try to stimulate that students would grasp all the concepts and their relations.

Tutors received training in handling the tutor instructions. It was stressed to only ask questions to keep the students actively engaged and not to explain the issues that were being discussed. Moreover, tutors were instructed to ask a question only when students were in danger of wandering off, when omissions were made, or when students did not know how to proceed. Finally, they were instructed to specifically emphasise the relations between the discussed concepts. For example, they were supposed to ask how sample size is related to the standard error.

At the end of the course the students evaluated the course on various aspects in a questionnaire. All courses are evaluated as a standard procedure. This is why students are used to filling out evaluation questionnaires after the final course exam. The questionnaire used in this study had 19 items consisting of statements that students had to rate on a five point Likert-scale. The statements covered the course itself, the statistical problems used in the meetings, the tutorial meetings, the performance of the tutor in stimulating understanding, and three more traditional aspects of the general functioning of the tutor, with respect to facilitation of the group and learning processes. The statements are presented in Appendix C. To assess students' achievement, the final course exam was used. This exam consisted of 30 multiple choice questions about the statistical and methodological subject matter of the course.

Analysis

The raw scores of the questionnaire and the final course exam scores were used for the analysis. The item scores were inspected with respect to the mode, skewness and kurtosis. A factor analysis was done on the individual items to distinguish the subscales. This resulted in five subscales, making up five dependent variables. Cronbach's alpha was computed for each variable. This study had a hierarchical design. First, the students were randomly assigned to the tutorial groups. Next, tutors were randomly assigned to the

guided and the control condition of this study. Because of this hierarchical design, the comparison between the two conditions, with respect to the five dependent variables, was done by means of multi-level analyses. Random intercept models were used for all five analyses, with the students as the first level and the tutors as the second level. Deviance tests were used for the random effects, because of the rather small sample size (Snijders and Bosker 2003).

Results

The evaluation questionnaire

Inspection of the item scores with respect to the mode, skewness and kurtosis showed no indication of violation of the normality assumption. Factor analysis resulted in two possible solutions. The scree plot indicated a three-factor solution (see Fig. 1).

The “eigenvalues greater than one criterion” indicated five factors. The eigenvalues for the five factors ranged from 7.00 to 1.02. We have opted for the five-factor solution, because after oblique rotation a clear pattern emerged that could directly be interpreted in relation to the content of the resulting five subscales (see Appendix C). The pattern matrix after oblique rotation is shown in Table 1. The highest factor loadings are in bold type. Items that related to the course itself loaded highly on the first factor (items 1–3). The sum score of these three items comprises the variable *course*. The first two items that related to the problems (items 4–5) loaded highly on the second factor; their sum comprises the variable *problems*. The other two items about the problems (items 6–7; referring to what the problems added to the discussions) together with three questions with respect to the discussions in the tutorial meetings (items 8, 10–11) loaded highly on the third factor. One item referring to these discussions (item 9) loaded equally on three different factors. Based on the content of the item it was categorised under the third factor. The sum of these six items (items 6–11) comprises the variable *elaboration*. Five items loaded highly on the fourth factor (items 12–16). These items were related to those aspects of the tutor functioning that were supposed to stimulate students’ understanding of the subject matter. The sum of these five items forms the variable *tutor guidance*. Finally, the three items with respect to the general functioning of the tutor (items 17–19) loaded highly on the fifth factor. Their sum forms the variable *tutor general*. Together the five factors explained

Fig. 1 Scree plot of the factor analysis on the individual items

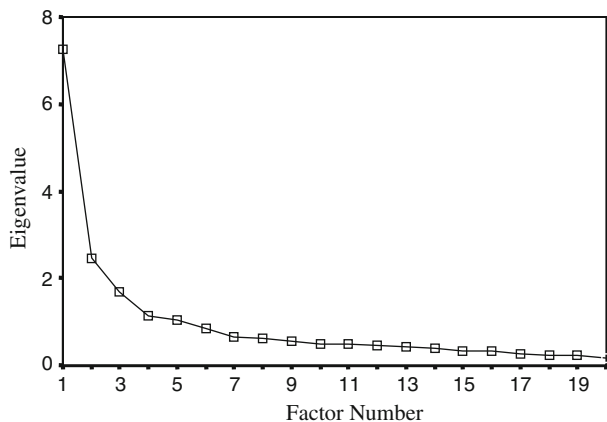


Table 1 Results of factor analysis on the items of the questionnaire

	Factor				
	Course	Problems	Elaboration	Tutor guidance	Tutor general
1. Clarity	.725	.048	-.100	-.009	-.101
2. Instructiveness	.779	-.002	.147	.007	.078
3. Organisation	.560	.010	-.030	.010	-.073
4. Variation	-.003	.737	-.050	-.087	.074
5. Systematic approach	.103	.404	-.128	.025	.128
6. Useful learning goals	.041	.193	-.613	.012	-.141
7. Useful discussions	.083	-.026	-.613	-.033	.175
8. Effectiveness	-.041	.105	-.777	.031	-.021
9. Integration	.105	.105	-.297	-.304	.301
10. Pleasant	.250	-.203	-.442	-.089	.311
11. Productive	.097	-.124	-.522	-.185	.319
12. Tutor systematically	-.002	-.014	-.187	-.596	.256
13. Tutor insight	-.035	-.022	-.118	-.754	-.047
14. Tutor structuring	.043	-.037	-.117	-.565	.264
15. Expertise use	-.018	.037	.164	-1.001*	-.131
16. Contributions relevant	.065	.058	.095	-.851	.045
17. Autonomy stimulation	.059	.016	-.148	-.248	.540
18. Activation	-.006	.042	-.183	-.133	.519
19. Evaluation	-.052	.161	.087	.021	.603
<i>Factor correlations</i>					
Course	1.000	.103	-.389	-.230	.165
Problems		1.000	-.291	-.087	.212
Elaboration			1.000	.317	-.409
Tutor guidance				1.000	-.482
Tutor general					1.000
Cronbach's α	.72	.60	.85	.90	.73

Note: Principal axis factoring with oblimin rotation

* Value larger than one, artefact due to oblique rotation

almost 68% of the variance. Cronbach's α for the five variables was high: *course* ($\alpha = .72$), *problems* ($\alpha = .60$), *elaboration* ($\alpha = .85$), *tutor guidance* ($\alpha = .90$), *tutor general* ($\alpha = .73$).

Fixed and random effects on the five dependent variables

The results of the multilevel analyses showed that in the guided condition the means of *course*, *elaboration*, and *tutor guidance* were significantly higher than in the control condition. No differences were found regarding *problems* and *tutor general*. The difference at the tutor level was not significant in the analysis of *course*. In the other four analyses these differences were significant. The results are presented in Table 2.

The means of the five variables for the two conditions are presented in Table 3.

Table 2 Random intercept multi level analyses on the five dependent variables with tutor as the second level

	<i>b</i>	S.E.	<i>t</i>	<i>p</i> -value		<i>p</i> -value
Fixed effects				Random effects		
<i>Course</i>						
Intercept	10.32	.63	16.29	.000*	τ^2_{tutor}	.11
Condition	-.99	.40	-2.48	.028*	Δ Deviance	.37 .270
<i>Elaboration</i>						
Intercept	20.83	1.45	14.36	.000*	τ^2_{tutor}	1.95
Condition	-2.65	.89	-2.97	.012*	Δ Deviance	14.40 .000*
<i>Tutor guidance</i>						
Intercept	23.99	1.97	12.17	.000*	τ^2_{tutor}	4.10
Condition	-4.11	1.21	-3.40	.006*	Δ Deviance	25.70 .000*
<i>Problems</i>						
Intercept	4.64	.02	7.19	.000*	τ^2_{tutor}	.35
Condition	.02	.38	.052	.959	Δ Deviance	13.60 .000*
<i>Tutor general</i>						
Intercept	9.75	1.08	9.06	.000*	τ^2_{tutor}	1.13
Condition	-1.06	.66	-1.60	.138	Δ Deviance	15.00 .000*

Note: * $p < .05$; Δ Deviance is the difference between the deviance of the model with and without the random tutor effect

Table 3 Means of the five dependent variables for the guided and control condition

	Number of items	Minimum	Maximum	Scale centre	M_{control}	M_{guided}
Course	3	3	15	9	8.4	9.3
Elaboration	6	6	30	18	15.6	18.3
Tutor guidance	5	5	25	15	16.0	19.9
Problems	2	2	10	6	3.5	4.1
Tutor general	3	3	15	9	7.6	8.7

The mean grade on the final course exam was higher for the students in the guided condition as hypothesised, although the difference was only marginally significant ($M_{\text{control}} = 5.85$, $M_{\text{guided}} = 6.20$; $t = 1.47$; $p = .072$).

Discussion

In this study, it was examined whether directive tutor guidance in problem-based learning of statistics improved the subjective perceptions of the students regarding the course, the tutor, the discussions in the tutorial meetings, the quality of the problems used in the meetings, and general tutor functioning. Directive tutor guidance aimed at stimulating the students to link together the statistical concepts in a structured way. It was expected that this would increase students' understanding of the topics and that students would gain a better insight into the content of the course.

The final course exam served as a measurement of achievement. The grades of the exam were used as an indication of students' understanding. We expected all students to be able to pass the exam at the end of the course, but we expected the students in the guided condition to do better. The results confirmed our hypotheses. We found a marginally significant effect. Moreover, students pass the exam if they get a grade of 5.5 (out of 10 points). The means of both conditions were above this cut off score. It should be noted, however, that the difference of 1.5 times the standard error in favour of the guided condition is very close to this cut off score. In the guided condition 67% of the students passed the final course exam, while 60% passed in the control condition. These proportions reveal that for a small group of students the positive effect of our intervention has been crucial. This means that for the individual student, the difference between the conditions may be quite relevant in practice.

With regard to the questionnaire the results showed that most hypothesised effects occurred. The course was valued more positively by the students in the guided condition. These students also rated the discussions in the tutorial meetings higher. Those aspects of the tutor functioning that were supposed to stimulate students' understanding of the subject matter were also better evaluated in the guided condition. In both conditions the problems were judged of equal quality. Furthermore, general functioning of the tutors was judged as similar in both conditions.

The items of the questionnaire referring to the course consisted of statements regarding the clearness of the goal of the course, the instructiveness of the course, and the organisation of the course. Students in the guided condition evaluated the course more positively than the students in the control condition. The more positive evaluation indicates that the students in the guided condition had better insight in the content of the course and had a more positive overall impression, as hypothesised. This result suggests that students apparently understood better what was expected from them with respect to the objectives of the course, what and how they had to learn, and what they did learn was clearer to them. All statements in the evaluation questionnaire had to be rated by the students on a five point Likert-scale. This means that a neutral opinion with respect to the course would have resulted in a mean score of nine, i.e. nine is the centre of the scale. The means of the two conditions show that students in the guided condition were slightly above this centre of the scale (i.e. a positive rating), where students in the control condition rated the course slightly negative.

Items constituting the variable elaboration included statements concerning the integration of the subject matter in the discussions and the productiveness of the meetings. Items referring to particularly those aspects of tutor behaviour that were supposed to stimulate students' understanding constituted the variable tutor guidance. These items included statements like: the tutor helped structuring the subject matter and the tutor's contributions were relevant. The higher ratings of the students in the guided condition of both elaboration and tutor guidance indicate that a more directive tutor behaviour also had a positive effect on the instructiveness of the discussions. A neutral stance toward elaboration would have resulted in a mean score of 18. The means for the variable elaboration show that students in the guided condition rated the discussions slightly positive, the students in the control condition negatively. Tutor guidance was judged positively in both conditions, as students in both conditions rated tutor guidance above 15, the centre of the scale.

Directive guidance in education may not only have positive effects, but it may also lead to a decrease of self-study time, autonomy and self-regulated learning, so students may

become less motivated and more passive (Ames 1992; Deci and Ryan 1985; Lepper et al. 1997; Pintrich 2003; Schmidt and Moust 2000). Therefore, directive tutoring might have led to a more negative evaluation of the category of items referring to the general functioning of the tutors. These items consisted of statements regarding the stimulation of autonomy, the activation of the students, and the evaluation of the group processes. However, no differences were found between the two conditions. It can be concluded that directive tutor guidance did not have a negative effect in this respect.

In both conditions the same problems were used to initiate the discussion in the tutorial group meetings. Therefore, no differences between the two conditions were hypothesised in students' judgments about the problems. This is exactly what we found. These findings together with the equal evaluation of the general functioning of the tutor in both conditions supported the other findings.

The results also show (except for the variable course) that tutors differ significantly from each other. The influence of the tutors on students' perception of the courses is relatively small. Their influence on the discussions and their own functioning is obviously much bigger. Multi-level analyses of those variables that are influenced by the tutors showed differences between tutors. For the course no differences were found. These results may seem trivial. However, for the uniformity of education it is important to try to increase similarity in the way tutors interact with the students. Our intervention exactly tried to do this. Tutors received specific questions and received training in how to use the instructions. This may have reduced the individual differences in tutoring style. As we did not directly observe tutor behaviour in the groups, but inferred this from students' responses on the evaluation questionnaire we do not know how the instructions were carried out in practice. Tutors who guide the discussion, as in our intervention, are more prominent in the meetings. As a consequence, directive guiding tutors might be inclined to explain some of the topics, although they were specifically instructed not to do so. Future research could be aimed at how the instructions influence differences between tutors and tutor behaviour *per se* in practice.

Future research could also be directed at the underlying mechanisms for the results that we found. Our results are in line with cognitive load theory. Providing guidance may have reduced extraneous cognitive load (van Merriënboer and Sweller 2005; Paas et al. 2004). Asking specifically for explanations of the relations between the statistical concepts may have increased germane cognitive load (Paas and van Merriënboer 1994; Sweller et al. 1998). It could be measured whether our approach has such an effect on the perceived cognitive load.

Finally, it is unclear which kind of students have profited most from the extra guidance. We assume that specifically poor students, who have had difficulty in mastering the subject matter, may have had the most benefit from the extra guidance. Future research is needed to confirm this.

This study was done in a field setting. On the one hand this limits the scope of our conclusions. We used a standard questionnaire and the regular final course exam as measurements. Moreover, we could not control all factors. For example, we do not have a complete view of how tutors behaved in the meetings, nor do we know how students studied in the different conditions. On the other hand, because the studied conditions were embedded in realistic learning situations, we think that the outcomes are relevant in practice.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

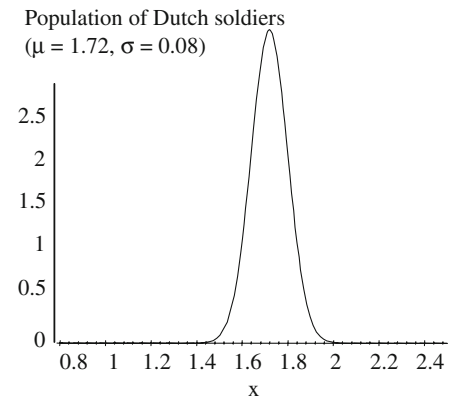
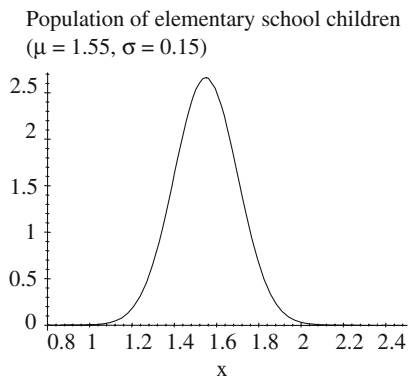
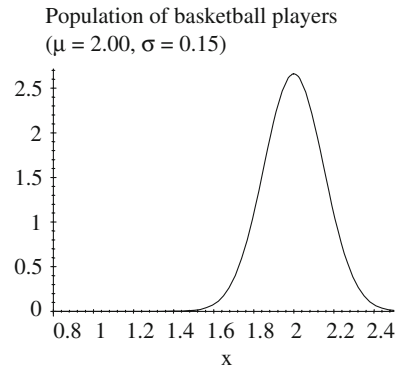
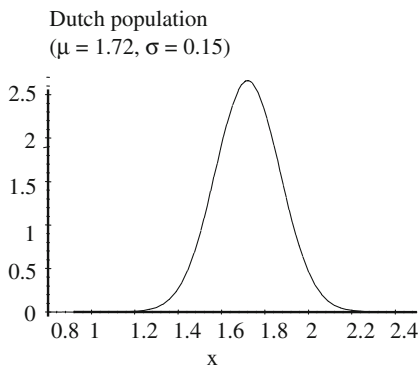
Appendix A

Example of a PBL problem for statistics

The presented problem is an introduction to the idea of sampling distributions. The problem tries to raise the notion of the three aspects (population mean μ , population standard deviation σ , and sample size N) that constrain the means of samples from a population. The problem is the same for both conditions. The ideas that students in the discussion intuitively come up with are elaborated and formalised in subsequent problems.

Problem 1. Expecting but not pregnant

Consider the following distributions of length:



- If someone was going to draw samples of 100 persons from the Dutch population, what do you think the *expected* value would be?
- What would the *expected* value be of the means of samples from the populations of basketball players and elementary school children?
- Now suppose that we draw smaller samples (for example of only three persons) from these populations. What values for the means of these samples can be found?

Let's look at the fourth distribution of length.

- Suppose we draw samples of 100 persons from the population of Dutch soldiers. Which values of the means of these samples are less likely as compared to the means of the samples from the Dutch population?

Can you indicate which 3 aspects determine your expectation of the mean of a sample that you have drawn?

Appendix B

Tutoring instruction of problem 1

The tutors in both conditions were given some explanations about what the problem is supposed to evoke, what the core topics are, and some clarification of the subject matter.

In addition, the tutors in the guided condition received the following guiding questions.

Example questions for the discussion (in the original frame the correct answers are also given).

- *What is the problem about? Why does the problem have this title? Which question is asked several times?*
- *What are the values that can be expected? How high is the probability of a sample mean (\bar{X}) of 2.85? Can the means have any value? Look at the three populations.*
- *Do you expect exactly the population mean in a sample?*
- *When will \bar{X} be higher than μ and when will it be smaller?*
- *Do you think that $\bar{X} = 2.10$ is very likely from a sample of $N = 100$ from the Dutch population?*
- *Why is that unlikely?*
- *Could it be the \bar{X} of a sample from the basketball players?*
- *Why is that?*
- *Give me ten values of people in a sample of $N = 10$ from the population of Dutch soldiers. What do you expect for \bar{X} ?*
- *Now calculate \bar{X} . Is it exactly μ ? Now substitute one of these ten values by 1.40, what happens to \bar{X} ?*
- *Was it possible to have a person of 1.40 in the sample?*
- *There are such small people in the Dutch population. What is your conclusion?*
- *Would it have had the same effect if we had substituted one of the values in a sample of $n = 100$? What is your conclusion?*
- *What can \bar{X} be from a sample of the Dutch population if $N = 1$; or $N = 3$; or $N = 10$?*
- *Which three aspects determine which \bar{X} 's are likely?*
- *How are distributions of possible \bar{X} 's called?*
- *Which theorem constraints the possible \bar{X} 's?*

Appendix C

Evaluation questionnaire

Statements of the questionnaire with respect to respectively the course, the problems, the tutorial meetings, tutors stimulating understanding, and general functioning of the tutors.

The course

1. The goal of the course has become clear to me
2. The course was instructive
3. The course was well organised

The problems

4. The problems showed sufficient variation
5. The problems evoked a systematic approach

Elaboration

6. The problems evoked useful learning goals
7. The problems evoked useful discussions
8. The discussions in the meetings were effective
9. In the discussions the subject matter was integrated
10. The meetings were pleasant
11. The meetings were productive

The tutor stimulating understanding

12. The tutor directed the discussions systematically
13. The tutor showed insight in the content of the course
14. The tutor helped structuring the subject matter
15. The tutors used their expertise to guide the group
16. The contributions of the tutors were relevant

Aspects of the traditional role of the tutor

17. The tutor stimulated autonomy
18. The tutor activated the students
19. The tutor regularly evaluated group processes

Note: These statements were rated by students on a five point Likert-scale

References

- Acker, S., Hill, T., & Black, E. (1994). Thesis supervision in the social sciences: Managed or negotiated? *Higher Education*, 28, 438–498.
- Albanese, M. A., & Mitchell, S. (1993). Problem-based learning: A review of literature on its outcomes and implementation issues. *Academic Medicine*, 68, 52–81.
- Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology*, 84, 261–271.
- Anderson, J. R. (1983). *The architecture of cognition*. Cambridge: Harvard University Press.
- Barrows, H. S. (1986). A taxonomy of problem-based learning methods. *Medical Education*, 20, 481–486.
- Barrows, H. S. (1988). *The Tutorial Process*. Springfield: Southern Illinois University School of Medicine.
- Carroll, W. M. (1994). Using worked examples as an instructional support in the algebra classroom. *Journal of Educational Psychology*, 86, 360–367.
- Chi, M. T. H. (1996). Constructing self-explanations and scaffolded explanations in tutoring. *Applied Cognitive Psychology*, 10, S33–S49.
- Chi, M. T. H., Bassok, M., Lewis, M. W., Reiman, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, 13, 145–182.
- Chi, M. T., Siler, S. A., Jeong, H., Yamauchi, T., & Hausman, R. G. (2001). Learning from human tutoring. *Cognitive Science*, 25, 471–533.
- Debowski, S., Wood, R. E., & Bandura, A. (2001). Impact of guided exploration and enactive exploration on self-regulatory mechanisms and information acquisition through electronic search. *Journal of Applied Psychology*, 86, 1129–1141.
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behaviour*. New York: Plenum Press.

- De Grave, W. S., Dolmans, D. H. J. M., & Van Der Vleuten, C. P. M. (1999). Profiles of effective tutors in problem-based learning: Scaffolding student learning. *Medical Education*, *33*, 901–906.
- Dolmans, D. H. J. M., Gijssels, W. H., Moust, J. H. C., De Grave, W. S., Wolfhagen, I. H. A. P., & Van Der Vleuten, C. P. M. (2002). Trends in research on the tutor in problem-based learning: Conclusions and implications for educational practice and research. *Medical Teacher*, *24*, 173–180.
- Dolmans, D. H. J. M., & Wolfhagen, I. H. A. P. (2005). Complex interactions between tutor performance, tutorial group productivity, and the effectiveness of PBL Units as perceived by students. *Advances in Health Sciences Education*, *10*, 253–261.
- Fay, A. L., & Mayer, R. E. (1994). Benefits of teaching design skills before teaching LOGO computer programming: Evidence for syntax independent learning. *Journal of Educational Computing Research*, *13*, 187–210.
- Gal, I., & Garfield, J. (1997). Curricular goals and assessment challenges in statistics education. In I. Gal & J. Garfield (Eds.), *The assessment challenge in statistics education*. Amsterdam: IOS Press.
- Graesser, A. C., Bagget, W., & Williams, K. (1996). Question driven explanatory reasoning. *Applied Cognitive Psychology*, *10*, S17–S131.
- Graesser, A. C., Person, N. K., & Magliano, J. P. (1995). Collaborative dialogue patterns in naturalistic one-to-one tutoring. *Applied Cognitive Psychology*, *9*, 495–522.
- Hmelo, C. E., & Evensen, D. H. (2000). Introduction. Problem-based learning: Gaining insights on learning interactions through multiple methods of inquiry. In D. H. Evensen & C. E. Hmelo (Eds.), *Problem-based learning: A research perspective on learning interactions*. Mahwah: Lawrence Erlbaum Associates Publishers.
- Lee, M., & Thompson, A. (1997). Guided instruction in LOGO programming and the development of cognitive monitoring strategies among college students. *Journal of Educational Computing Research*, *16*, 125–144.
- Lepper, M. R., Drake, M. F., & O'Donnell-Johnson, T. (1997). Scaffolding techniques of expert human tutors. In K. Hogan & M. Pressley (Eds.), *Scaffolding student learning: Instructional approaches and issues*. Cambridge: Brookline Books.
- Paas, F. G. W. C. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*, *84*, 429–434.
- Paas, F., Renkl, A., & Sweller, J. (2004). Cognitive load theory: Instructional implications of the interaction between information structures and cognitive architecture. *Instructional Science*, *32*, 1–8.
- Paas, F., & van Merriënboer, J. J. G. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive load approach. *Journal of Educational Psychology*, *86*, 122–133.
- Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology*, *95*, 667–686.
- Schau, C., & Mattern, N. (1997). Assessing students' connected understanding of statistical relationships. In I. Gal & J. B. Garfield (Eds.), *The assessment challenge in statistics education*. Amsterdam: IOS Press.
- Schmidt, H. G., v. d. Arend, A., Moust, J. H. C., Kokx, I., & Boon, L. (1993). Influence of tutors' subject-matter expertise on student effort and achievement in problem-based learning. *Academic Medicine*, *68*, 784–791.
- Schmidt, H. G., & Moust, J. H. C. (1995). What makes a tutor effective? A structural-equations modelling approach to learning in problem-based curricula. *Academic Medicine*, *70*, 708–714.
- Schmidt, H. G., & Moust, J. H. C. (2000). Factors affecting small-group tutorial learning: A review of research. In D. H. Evensen & C. E. Hmelo (Eds.), *Problem-based learning: A research perspective on learning interactions*. Mahwah: Lawrence Erlbaum Associates Publishers.
- Smeby, J.-C. (1996). Disciplinary differences in university teaching. *Studies in Higher Education*, *21*, 69–79.
- Snijders, T. A. B., & Bosker, R. J. (2003). *Multilevel analysis. An introduction to basic and advanced multilevel modeling*. London: Sage Publications Ltd.
- Sweller, J. (1999). *Instructional design in technical areas*. Mahwah, NJ: Erlbaum.
- Sweller, J., & Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction*, *2*, 59–89.
- Sweller, J., van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*, 251–296.
- Tuovinen, J. E., Sweller, J. (1999). A comparison of cognitive load associated with discovery learning and worked examples. *Journal of Educational Psychology*, *91*, 334–341.
- van Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review*, *17*, 147–177.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.