

## Is learning in problem-based learning cumulative?

Elaine H. J. Yew · Esther Chng · Henk G. Schmidt

Received: 31 August 2010 / Accepted: 3 December 2010 / Published online: 16 December 2010  
© The Author(s) 2010. This article is published with open access at Springerlink.com

**Abstract** Problem-based learning (PBL) is generally organized in three phases, involving collaborative and self-directed learning processes. The hypothesis tested here is whether learning in the different phases of PBL is cumulative, with learning in each phase depending on that of the previous phase. The scientific concepts recalled by 218 students at the end of each PBL phase were used to estimate the extent of students' learning. The data were then analyzed using structural equation modeling. Results show that our hypothesized model fits the data well. Alternative hypotheses according to which achievement is predicted either by collaborative learning alone or by self-directed learning alone did not fit the data. We conclude that the learning in each PBL phase is cumulative, and strongly influenced by the earlier phase, thus providing support for the PBL cycle of problem analysis, self-directed learning, and a subsequent reporting phase. We also demonstrate an efficient method to capture and quantify students' learning during the PBL process.

**Keywords** Problem-based learning · Learning process · Collaborative learning · Self-directed learning

### Introduction

Educators have long been advocating 'active' learning whereby students are engaged in meaningful activities as part of their learning process. Active learning has been generally defined as any instructional strategy that involves "students in doing things and thinking about what they are doing" (Bonwell and Eison 1991, p. 2). Given such a broad definition, active learning can be viewed as encompassing a wide variety of instructional methods.

---

E. H. J. Yew (✉) · E. Chng  
Centre for Educational Development, Republic Polytechnic, 9 Woodlands Avenue 9,  
Singapore 738964, Singapore  
e-mail: Elaine\_yew@rp.sg

H. G. Schmidt  
Department of Psychology, Erasmus University, P.O. Box 1738,  
3000 DR Rotterdam, The Netherlands

Although various studies have demonstrated the effectiveness of promoting student engagement using interactive-engagement methods compared to those in traditional courses (reviewed by Michael 2006; Prince 2004), questions about *how* students learn while being actively engaged, both individually and when in collaborative small groups, remain to be further investigated.

Generally learning is thought to be a cumulative process where new learning builds upon knowledge acquired in a previous phase. In the case of active learning, it is assumed that both collaborative learning episodes and individual self-directed study phases play important roles in students' learning. Although the idea that new learning is dependent on what has been learned previously is almost universally accepted, demonstrations of its truth have been largely confined to the psychological laboratory, particularly in the field of text processing (e.g., Bransford and Johnson 1972; Kintsch and Van Dijk 1978). To our knowledge, no natural classroom demonstration of the cumulative nature of learning exists to date. Moreover, since social constructivism suggests that knowledge is mainly constructed by means of collaborative interactions (e.g., Cobb 1994; Driver et al. 1994), it is possible that the effects of active learning on achievement are really only due to the group interactions and co-construction of knowledge. Alternatively, since research on self-regulated learning has shown that the use of self-regulated learning strategies strongly influences academic achievement (Zimmerman 1990), it can be argued that it is the individual self-directed learning phase that is most important to students' learning.

The purpose of this paper therefore is find the extent to which active learning is cumulative and whether it involves both collaborative and self-directed learning, in the context of problem-based learning (PBL). PBL is an example of an active-learning approach in which students are given the opportunity to learn independently as well as collaboratively, while understanding an ill-structured problem. It was originally developed in medical schools to help students integrate basic science and clinical knowledge, as well as to develop clinical reasoning and lifelong learning skills (Barrows 1986). However it is now of increasing interest to educators of various levels and disciplines (Gallagher et al. 1992; Kolodner et al. 2003) as it provides a structured framework of active and collaborative learning, in line with current understanding of learning as a constructive and co-constructive activity involving social interactions (Glaser and Bassok 1989; Palincsar 1998). As will be described in greater detail later on, PBL involves a sequential series of learning phases that emphasizes collaborative and individual self-directed learning at different points in time. The assumption underlying PBL is that learning in the PBL process is cumulative—learning in one phase is dependent on the previous, and also that both co-construction with peers and individual construction of concepts during self-directed study contribute to student learning (Schmidt 1983). We therefore seek to test the assumptions regarding the nature of learning in PBL, by tracing the learning process of students throughout all the phases of PBL. The central thesis to be tested is whether learning in the different phases of PBL is cumulative—does the learning in each phase depend on the previous phase? Or are some phases of the PBL process more (or less) important than others? Secondly, we also seek to understand how students learn in the different phases of PBL in terms of concept acquisition and elaboration. A third objective is to devise an efficient and valid method to track students' learning as it unfolds in the course of the PBL process.

## The PBL process

PBL always starts with a problem, for which students do not prepare beforehand. After the description of the problem is given to small groups of students, they first analyze the

problem, generate possible explanatory hypotheses, build on one another's ideas, as well as identify key issues to be studied further. These activities allow students to construct a shared initial explanatory theory or model explaining the problem-at-hand based on their prior knowledge (Schmidt 1983). After this period of teamwork, they disperse for a period of individual study to work on learning issues they have identified as a group. When they next meet as a team during what is called the "reporting phase", they are expected to share and discuss their findings, as well as refine their initial explanations based on what they have learned. Students would then move on to analyze a new problem, or if new learning issues requiring further study are identified during this phase, the process described above would be repeated. Thus, PBL can be seen as a cyclical process consisting of three phases: initial problem analysis, self-directed individual learning, and a subsequent reporting phase (Barrows 1988; Hmelo-Silver 2004; Schmidt et al. 2009). A tutor is present to guide students' learning in the problem analysis and reporting phases. The tutor's role is to facilitate the processes involved when students co-construct knowledge through discussions and sharing of ideas (Hmelo-Silver and Barrows 2006). In PBL, both group and individual learning processes are recognized to play important supplementary roles in students' learning (Schmidt and Moust 2000; Van den Hurk et al. 2001).

#### How students learn in the process of PBL

Various studies have focused on how students learn in the different phases of the PBL cycle. The initial problem analysis activates students' prior knowledge and allows them to relate new information in the problem to their existing knowledge. Hearing what other students elaborate upon could also serve to activate or uncover the less accessible prior knowledge in the listeners. Studies by De Grave et al. (2001) and Schmidt et al. (1989) have demonstrated that elaboration during problem analysis in a small group prior to studying problem-relevant new information resulted in increased knowledge acquisition and recall. As argued by De Grave et al. (1996), such elaboration and activation of existing knowledge are instrumental in restructuring and transferring concepts resulting in the construction of new knowledge and ideas. The process of discussion during the problem analysis phase would also result in students realizing the gaps between their existing knowledge and what they are required to know in order to respond to the problem. Thus students would identify these gaps as learning issues to be studied further during the self-directed learning phase. This individual study phase is a key feature in PBL, in line with its underlying "student-centered" philosophy of enabling students to take responsibility for their own learning by deciding what to study and to what extent. Through the self-directed learning phase, students learn important skills such as goal setting, planning and self-control in terms of time and task-management (Zimmerman 2002). As students implement their course of actions to achieve their goals, they would also have to monitor and reflect on their own progress, thus exhibiting a kind of feedback loop in the process (Hmelo-Silver 2004; Zimmerman 1990).

When the tutorial group reconvenes to report their findings and the results of their individual study, opportunities are given to students to present, explain and defend their ideas, and in the process, to restructure or refine their own knowledge networks (Schmidt and Moust 2000). The discussions during the reporting phase are centered on students' response to the problem statement given in the problem analysis phase. Studies have shown that group interactions such as elaborations and co-constructions take place during this phase, allowing for collaborative knowledge construction (Hmelo-Silver and Barrows 2008; Visschers-Pleijers et al. 2004; Yew and Schmidt 2008).

## Factors influencing students' learning in PBL

A few studies have examined and tested how the variables thought to be active in PBL influence and relate with one another and students' learning outcomes. Gijsselaers and Schmidt (1990) tested a path model relating input variables such as the quality of problems, tutor performance and students' existing knowledge, process variables such as group functioning and time spent on self-directed study, and the outcomes of learning. They demonstrated that problem quality influences tutorial group functioning, which in turn had an influence on the amount of time spent in individual study. More time put into individual study led to increased academic achievement. This model was further refined by Van den Hurk et al. (2001). They investigated in more detail what actually happens to learners during problem analysis, individual study and reporting. They found that the quality of learning issues generated during the problem analysis phase had an impact on the extent to which the learning issues were used during individual study. Increased usage of learning issues during self-directed study also influenced students' research to be more explanation-oriented, which in turn led to a "deeper discussion" during the reporting phase. Finally the depth of reporting led to a higher score on an achievement test.

Both of these tests of a causal model provide insight into the relationships between the variables important in the PBL process and hence into how students learn in PBL. In particular, the study by Van der Hurk et al. suggests that learning in PBL is indeed cumulative. Their study demonstrates that learning in the problem analysis phase influences individual study, which in turn influences the reporting phase, and finally achievement. However, as recognized by the authors, a limitation to both studies was that data were obtained based on students' perceptions and retrospective self-report rather than on their actual behaviors. As argued by Dolmans and Schmidt (2006), and Hak and Maguire (2000), the research required to uncover the relationships between aspects of the tutorial process and students' learning should be focused on the *actual* activities occurring in the various phases of PBL.

Some studies have used direct observational methods to examine how and what students learn during PBL. One observational study focusing on the content of the learning-oriented interactions of students was conducted by Yew and Schmidt (2008). Here the verbal interactions taking place in an entire PBL process were audio recorded and analyzed qualitatively. While the results demonstrated that PBL stimulates constructive, self-directed and collaborative learning processes, no relationships between the content of their interactions with subsequent learning were reported. In addition, due to the data- and time-intensive nature of the methodology involved, the sample size used in the study was limited, thus making statistical analysis difficult. A recent study by Hmelo-Silver and Barrows (2008) analyzed in detail the knowledge building process in a PBL tutorial by examining the discourse of students and facilitator throughout both the problem analysis and reporting phase of a PBL tutorial. This was carried out by videotaping five students as they worked on a problem for more than 5 h in two separate sessions. The study demonstrated how an expert facilitator guided the group discourse with the use of open-ended metacognitive questions, and how students actively worked on enhancing and refining their collective knowledge throughout the group interaction portions of a PBL cycle. However this study again did not relate the quality of students' verbal contributions to outcomes of their learning.

There have been several other attempts to trace the learning process in PBL. Visschers-Pleijers et al. (2006) made use of video recording while other researchers have made use of stimulated recall (De Grave et al. 1996), and thought sampling (Geerligs 1995) to provide

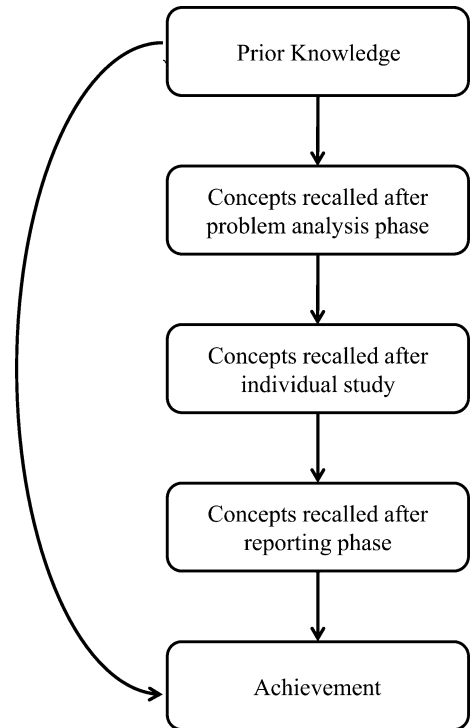
qualitative descriptions of the actual behaviors and activities in a PBL tutorial. The difficulty of such approaches is that they do not easily allow for the quantification of learning. In addition, they are so data-intensive that studying larger numbers of students becomes almost impossible. A case in point is our own previous attempt to study and identify the relationships between learning activities of students in PBL with their learning outcomes (Yew and Schmidt 2008). We recorded all verbal interactions of two groups of students for an entire PBL cycle. In addition, we logged all their individual study activities, which were conducted through the use of computers. The resulting protocols, consisting of around 72 h of material were segmented into 'idea units' consisting of the scientific ideas that were exchanged and studied (Meyer 1985). The units of analysis selected were the relevant scientific concepts found in the protocols as expressed by the individual students during discussion and encountered during individual study on the internet (more about the relevance of scientific concepts for studying learning online in [Method](#) section). We identified and counted the relevant scientific concepts articulated by each student during the different PBL phases and those they studied individually while working on the problem-at-hand. By analyzing the number of concepts acquired over the different learning phases for the nine students two distinct phases in the PBL process were identified—an initial concept articulation phase, in which students are exposed to and articulate new ideas, and a later concept repetition phase, in which ideas acquired seem to be repeated and elaborated upon. Given the small number of students involved, however, further statistical analysis of the data proved impossible. A second study using the same methodology included a larger sample size of 35 students and thus enabled us to analyze the quantitative relationships between students' verbal interactions during different phases of the problem-based learning cycle, self-directed study, and achievement, using a structural equation modeling approach (Yew and Schmidt 2010). Our results showed that students' verbal contributions through collaborative discussion during the initial problem analysis phase strongly influenced the extent of their verbal contributions in the reporting phase. Greater contribution of relevant concepts verbalized during the reporting phase also led to higher achievement at the end of the PBL cycle.

The methodology as used in these studies assumes that exposure to (from computer screen recordings of internet study resources) or the articulation of a concept during discussion can be considered a proxy of the learning taking place. However it is possible that students may not really understand the concepts they were verbalizing, or could be simply scanning the computer screens without seriously studying the material before them. In addition, the recording and transcription of all the learning activities throughout a PBL cycle turned out to be extremely time-consuming, thus limiting the sample size that could be utilized for each study.

To our knowledge, no natural classroom demonstration of the cumulative nature of learning exists to date. Moreover, since social constructivism suggests that knowledge is mainly constructed by means of collaborative interactions (e.g. Cobb 1994; Driver et al. 1994), it is possible that the effects of active learning on achievement are really only due to the group interactions and co-construction of knowledge. Alternatively, since research on self-regulated learning has shown that the use of self-regulated learning strategies strongly influences academic achievement (Zimmerman 1990), it can be argued that it is the individual self-directed learning phase that is most important to students' learning.

The purpose of this paper therefore is find the extent to which active learning is cumulative and whether it involves both collaborative and self-directed learning, in the context of problem-based learning (PBL). [Figure 1](#) summarizes our hypothesized relations in terms of a causal model. We hypothesized that learning in PBL is a cumulative process

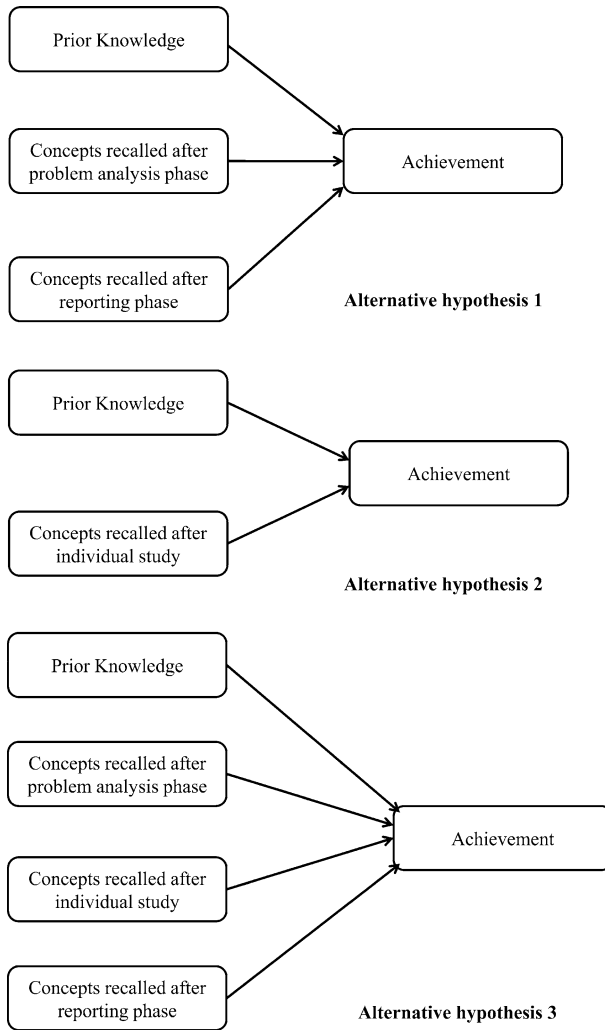
**Fig. 1** Hypothesized model on the relationships between the different learning phases of PBL



where the learning in each new phase builds upon knowledge acquired in a previous phase. The process is initially driven by the prior knowledge that students bring with them to the classroom and the learning in each of the PBL phases influences student achievement.

As mentioned earlier, it could be argued that the effects of active learning on achievement are mainly due to the group interactions and co-construction of knowledge or alternatively, that it is the individual self-directed learning phase that is most important to students' learning. We therefore test our hypothesis against these alternative hypotheses: (1) Learning in PBL is only influenced by phases involving collaborative learning and co-construction; (2) Learning in PBL is only influenced by self-directed study; and (3) Learning in PBL is influenced by both collaborative learning as well as self-directed study, but not in a sequential cumulative manner. These alternative models are summarized in Fig. 2.

Secondly, we hypothesize that the different PBL phases would involve the acquisition of new ideas (concepts) and the elaboration of previously acquired concepts to different extents. In an earlier preliminary study involving only nine students, we have shown that two different phases of the PBL process could be observed: an initial *terminology articulation* phase—consisting mainly of the problem analysis phase and initial SDL period, and characterized by the emergence of new concepts articulated and studied online, and secondly, a *terminology repetition* phase (mainly the later part of the SDL phase) where relevant concepts are repeated (Yew and Schmidt 2008). Here we aim to test this “acquisition-elaboration theory” of learning in PBL again, this time using a larger sample size. Finally, an important auxiliary issue is: How can students' learning be recorded as it unfolds? Through this study, we also aimed to develop and evaluate an efficient method to



**Fig. 2** Alternative models on the relationships between the different learning phases of PBL

capture and quantify students’ learning during the PBL process so that causal relationships in the PBL process can be identified through path analysis.

**Method**

Participants

Participants were 218 students from 11 randomly selected classes. The students were in their second year in the School of Applied Science at a polytechnic in Singapore. Data were collected from these students during the third week of their Molecular Cell Biology class. As they had already completed 1 year of study in the polytechnic, students were not new to the PBL approach described below. Students and facilitators gave informed consent.

## Educational context

The PBL process adopted at this polytechnic is somewhat unique in its “One-day-one-problem” approach. Here students work on one problem per day. Each class has a maximum of 25 students working together in teams of five. A brief description of the day’s process is described below:

- Problem analysis phase (approximately 1 h): The facilitator presents the problem for the day. Students work in teams of five to identify their prior knowledge and learning issues.
- Self-directed learning (SDL) period (approximately 4 h): Students do individual study or work with their teams on worksheets and other resources provided. They are also able to access other resources from the internet or textbooks. Time is spent helping one another within the team when necessary. Students meet with their facilitator for about 20 min in between this period to share their learning progress and strategy of understanding the problem.
- Reporting phase (approximately 2 h): Each team presents their consolidated findings and response to the problem, defending and elaborating based on questions raised by peers and the facilitator. The team presentation is usually in the form of powerpoint slides and the facilitator would also clarify key ideas if necessary.

Although the PBL process in this institution was adapted to suit the learning needs of the students and is completed within 1 day, it remains classified as PBL based on the “six core characteristics of PBL” described by Barrows (1996). The characteristics include student-centered learning whereby students work in small groups under the guidance of a tutor who facilitates the learning process. Problems are used as the stimulus for students’ learning with no opportunities to prepare beforehand. Furthermore, facilitators do not provide direct instruction. Instead, students construct their own understanding through self-directed learning (Hmelo-Silver 2004). An additional feature of the PBL approach in this context is that instead of only individual study during the self-directed learning phase, peer consultation and collaboration also takes place during this time.

## Procedure

A concept recall exercise was designed to estimate the number of relevant concepts that students were able to recall at the end of each PBL phase: problem analysis, self-directed learning and reporting. Our assumption is that as students engage in problem analysis, self-directed learning, group discussions, and/or peer teaching, they would be building networks of concepts related to the different learning issues as well as making relations between their prior knowledge and new ideas (Glaser and Bassok 1989). A beginner’s initial network would consist of a few isolated concepts or ideas that are poorly connected. Therefore, if asked to retrieve relevant concepts from these cognitive structures, his or her memory will be limited. The more students have learned about a topic, the richer, more coherent, and more detailed this particular network would be (Glaser and Bassok 1989). As learning progresses, more linkages and integration between new and existing ideas are constructed. Therefore, students who have learned more effectively would be able to recall more concepts and would do that more easily (Collins and Quillian 1969; Rumelhart and Norman 1978). Hence, measuring the number of relevant concepts students were able to recall in regards to the problem-at-hand at the end of each learning phase gives an



indication of the quality of students' learning, as well as the concepts they were exposed to either from what they had read or discussed during that phase.

The concept recall exercise was given to the students three times in the day—at the end of the problem analysis phase, self-directed learning and reporting phase. It consisted of the following instruction: “List all the keywords or terminologies that are related to DNA and/or RNA.” (Understanding the structure of DNA and RNA was the focus of the particular day's learning.) Students were instructed to only list concepts or keywords they thought were relevant, and not write in paragraphs or sentences. They were not allowed to discuss their answers or to refer to any resources when completing the exercise.

## Materials

The problem statement for the day was entitled “Made for the Job” and it introduced students to concepts related to the structures and functions of DNA and RNA. A week prior to the problem, students were given an essay pre-test consisting of the following instruction: “Describe and explain as much as you know about the *structure* of DNA and RNA.” This was to measure students' prior knowledge in regards to the topic. The same essay question was administered as a post-test immediately after the day's problem to measure students' learning achievement. No time limit was set but students were instructed to complete the test on their own without referring to any resources. The problem statement is presented in [Appendix](#).

The “idea unit” was used as the entity for scoring the free recall essay tests for accuracy (Meyer 1985; Schiefele and Krapp 1996). Answers were segmented into idea units, which was defined as a statement ending with a comma, period, or “and”. A score of 2, 1 or 0 was awarded to each idea unit. A score of 2 was given for a completely correct idea unit, 1 for a partially correct idea unit and 0 when the idea unit was completely incorrect. The first and second authors independently scored about 20% of the tests with inter-rater correlation of  $r = .91$ . The remaining tests were scored by the first author.

## Analysis

Students' answers to the concept recall procedure were analyzed by awarding 1 point to each relevant concept given by the student. These concepts (keywords and terminologies related to DNA and RNA) were agreed upon by the first and second authors before rating. Both authors have expertise in the field of molecular and cell biology. All the concepts were then rated by both authors and checked for differences. As the keywords and terminologies related to DNA and RNA were not ambiguous, they were scored with only one discussion between the two raters to establish consistency. Total scores of each student for each PBL phase were then further analyzed.

The relevant concepts listed were counted for each student for each learning phase (i.e. problem analysis, self-directed learning and reporting). The *total* number of concepts refers to the total number of relevant concepts recalled, including those that were repeated in one session. *Newly emerged* concepts were those that were not previously mentioned by the individual in any prior learning phase of the day. *Repeated* concepts were those that were previously recalled in an earlier learning phase. For the problem analysis phase, newly emerged and repeated concepts were deduced by comparing concepts listed at the end of the phase during the concept recall exercise with the concepts written in the pre-test answers.

*T* tests were used to compare differences in pre- and post-test results. One-way ANOVA was used to find out if there were significant differences in the mean number of relevant concepts recalled at the end of each learning phase. The data were also analyzed using structural equation modeling (SEM), a method that is able to test causal hypotheses among multivariate data. The pre- and post-test results as well as the total number of relevant concepts recalled by the students at the end of each PBL phase were analyzed for this structural equation modeling analysis. The method generates several statistics that enable the investigators to assess how well the empirical data fit the theoretical model and to estimate the strengths of the causal relations hypothesized. Four indicators suggested in the literature were used to evaluate the goodness-of-fit of the models to the sample data, namely, the Chi-square/*df* index of fit, Chi-square, the Comparative Fit Index (CFI), and the Root Mean Square Error of Approximation (RMSEA) (Arbuckle 2006; Browne and Cudeck 1993; Hu and Bentler 1999). The level of significance (*p*) computed from Chi-square and degrees of freedom should be higher than 0.05. The Chi-square/*df* index of fit yielded by dividing the minimum discrepancy (*C*) by its degrees of freedom should be lower than 3 and preferably close to 1 (Arbuckle 2006). CFI values larger than 0.95 and RMSEA scores below 0.06 can be considered as indicators of good fit (Browne and Cudeck 1993).

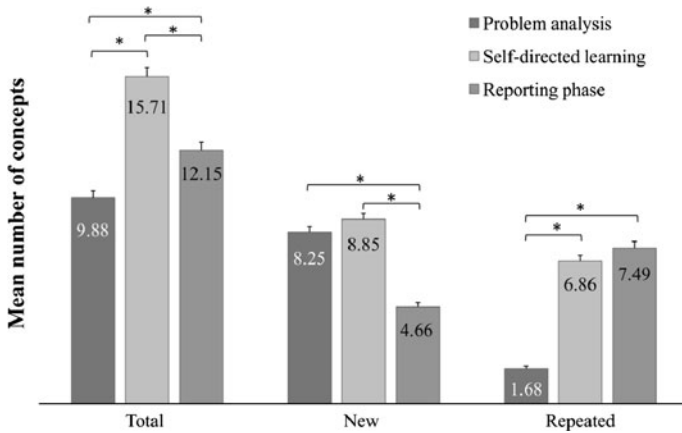
## Results

Results of mean student performance for the free recall essay pre- and post-tests showed improved scores for the post-test. The average difference between the post-test and pre-test scores for the free recall essay questions was 4.88 (*SD* = 3.88), indicating a significant increase in achievement at the end of the learning process,  $t(217) = 21.31, P < .01$ . The pre- and post-tests were significantly correlated at  $r = .44, P < .01$ .

The relevant concepts recalled by students at the end of each learning phase during the concept recall exercise were counted in three different ways—the total number of relevant concepts including those which were repeated, newly emerged concepts as well as repeated concepts. The distribution of the average number of these relevant concepts is shown in Fig. 3.

The one-way ANOVA revealed that the concepts verbalized differed significantly as a function of the different learning phases. The assumption of homogeneity of variance was violated so the Brown-Forsythe *F*-ratio is reported. There was a significant effect of the learning phase on the total number of concepts,  $F(2, 618.13) = 55.59, P < .01$ ; number of newly emerged concepts,  $F(2, 609.93) = 79.32, P < .01$  and repeated concepts,  $F(2, 497.73) = 156.06, P < .01$ .

Post-hoc analyses using the Games-Howell test showed that the total number of relevant concepts recalled was significantly higher after the self-directed learning phase ( $M = 9.88, SD = 4.86$ ) as compared to after the problem analysis phase ( $M = 15.71, SD = 6.52$ ) ( $P < .05$ ) and the reporting phase ( $M = 12.15, SD = 5.94$ ). The total number of relevant concepts recalled after the reporting phase was also significantly higher compared to after the problem analysis phase. For the number of newly emerged concepts, these were significantly higher in the problem analysis phase ( $M = 8.25, SD = 4.10$ ) and self-directed learning phase ( $M = 8.85, SD = 4.06$ ) compared to the reporting phase ( $M = 4.66, SD = 2.99$ ), while for the repeated concepts, these were significantly higher in the self-directed learning phase ( $M = 6.86, SD = 4.19$ ) and reporting phase ( $M = 7.49, SD = 4.71$ ) compared with the problem analysis phase ( $M = 1.68, SD = 1.85$ ). These significant differences are indicated in Fig. 3.



**Fig. 3** Distribution of the mean number (+SE) of total, new and repeated relevant concepts recalled at the end of the different learning phases of the PBL process ( $N = 218$ )

Table 1 shows the intercorrelations, means and standard deviations of the variables used in the structural equation model. Prior knowledge as measured by the essay pre-test is significantly correlated to students’ learning achievement and the total number of concepts recalled after each of the PBL phases. Students’ achievement is also significantly correlated to the concepts recalled at the end of each PBL phase. It can also be seen that the concepts recalled at the end of the different PBL phases are highly correlated with one another.

The hypothesized model displayed in Fig. 1 was tested against the data, yielding the following results: Chi-square = 7.84,  $df = 5$ ,  $P = .17$ ; the minimum discrepancy,  $C$ , divided by the degrees of freedom, Chi-square/ $df = 1.57$ ; the square root of the population discrepancy corrected by the complexity of the model  $RMSEA = .05$ ; and the Comparative Fit Index (CFI) = .97. Figure 4 displays the path diagram of the model, showing the significant paths. The parameter estimates for the model were all statistically significant. These findings show that the model fits the data adequately.

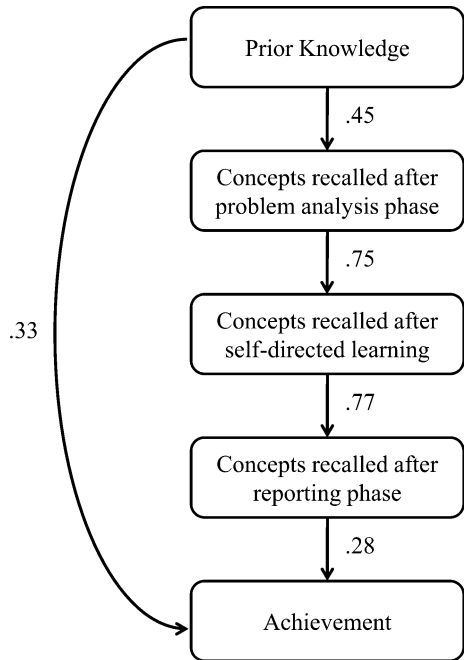
The following alternative hypotheses proposed in the introduction were also tested against the data: (1) Learning in PBL is only influenced by phases involving collaborative learning and co-construction; (2) Learning in PBL is only influenced by self-directed study; and (3) Learning in PBL is influenced by both collaborative learning as well as

**Table 1** Intercorrelations, means and standard errors of the variables ( $N = 218$ )

	1	2	3	4	5
1. Pre-test results (prior knowledge)	–				
2. Total number of concepts recalled after problem analysis	.44**	–			
3. Total number of concepts recalled after individual study	.37**	.74**	–		
4. Total number of concepts recalled after reporting phase	.26**	.59**	.75**	–	
5. Post-test results (achievement)	.41**	.34**	.40**	.34**	–
Mean	2.02	9.88	15.71	12.15	6.90
Standard deviation	2.05	4.86	6.52	5.94	3.66

\*\* Significant at the 0.01 level

**Fig. 4** Path model of the hypothesized model on relationships between different PBL phases



**Table 2** Comparison of results for different models tested using structural equation modeling

Model\indicators of goodness of fit	Cmin	df	Cmin/df	P value	CFI	RMSEA
Hypothesized model: learning in PBL is influenced by both collaborative learning and self-directed learning in a cumulative manner	7.84	5	1.57	.17	.97	.05
Alternative model 1: learning in PBL influenced only by collaborative learning	70.18	3	23.39	.00	.19	.32
Alternative model 2: learning in PBL influenced only by self-directed learning	26.33	1	26.33	.00	.41	.34
Alternative model 3: learning in PBL is influenced by both collaborative and self-directed but not in a cumulative manner	103.12	6	17.19	.00	.12	.27

self-directed study, but not in a cumulative manner. Table 2 shows a comparison of the results of the indicators of goodness of fit for the different models tested.

**Discussion**

Our goals in this study were to understand how students learn in the different phases of PBL in terms of concept acquisition and elaboration as well as to investigate the extent to which active learning is cumulative and whether it involves both collaborative and self-directed learning, in the context of PBL. In addition, we sought to devise an efficient and valid method to track students’ learning in the PBL process.

The results in Fig. 3 indicate that the self-directed learning phase is rich both in the acquisition of new concepts as well as the reiteration and repetition of concepts previously exposed to. That there was a high number of new concepts at the end of the problem analysis phase (as compared to students' pre-test answers) suggests that the discussion during this phase helped to activate students' prior knowledge, as previous studies have suggested (De Grave et al. 2001; Schmidt et al. 1989). We also observe that the reporting phase is characterized more by repetition of concepts rather than being exposed to new ones. This result is similar to the findings by Yew and Schmidt (2008) who identified two distinct phases of initial terminology articulation and a later terminology repetition in the PBL process from a group of students' online research data and verbal interactions. Our results strengthen their findings, which were limited due to small sample size.

One surprising observation from the distribution of concepts in Fig. 3 is that the total number of concepts recalled during the reporting phase is less than that in the self-directed learning phase. One would expect that by the end of the whole PBL cycle, students would be able to recall more relevant concepts. Possible reasons for this observation could be due to students already starting to forget some of the concepts learned within the day, or they could also be mentally drained by the end of an intensive day's work.

Our hypothesis in this study is that learning in PBL is a cumulative process where the learning in each new phase builds upon knowledge acquired in a previous phase. Results from Table 2 clearly show that compared with the alternative hypotheses, our hypothesized model best fits the data obtained. This model shows that there was significant impact of students' prior knowledge on the concepts students were able to recall after the problem analysis phase (.45). Students' prior knowledge also influenced their achievement directly (.33). This finding is in line with a previous study by Gijsselaers and Schmidt (1990) who found that amount of prior knowledge influenced students' achievement by .37. The number of relevant concepts recalled at the end of the problem analysis phase strongly influenced the number recalled at the end of the self-directed learning phase, which similarly influenced the number of concepts recalled at the end of the reporting phase. Finally being able to recall more relevant concepts at the end of the reporting phase influenced students' learning achievement significantly (.28). Results from the alternative hypotheses tested as tabulated in Table 2 also show that learning in PBL cannot be described only in terms of collaborative learning and teamwork, nor only in terms of self-directed learning. The lack of fit of the models with the data also demonstrates the importance of the sequential influence of learning from one phase to the next. This is important evidence showing that the three phases of PBL: problem analysis, self-directed learning, and reporting phase, play specific roles in influencing students' learning achievements.

Since our model enables us to predict student achievement very well, this also indicates the validity of our methodology as a means of keeping track of students' learning in the course of the learning process. Thus our method appears to be a useful and efficient way to overcome the typical difficulties faced in data collection of large samples for naturalistic studies.

One limitation of this present study is that the units of analysis focused on individual scientific concepts students were able to associate with the topic-at-hand and to recall at the end of each PBL phase, without connecting propositions demonstrating how the different concepts were linked. This then limits the deductions we can draw about the depth and accuracies of students' understanding of the different concepts. However despite this shortcoming, our findings from our model fit also show that this method does provide valid insight into students' learning.

In conclusion, we have shown that all the phases in the PBL process are necessary to understand how students learn in PBL. The learning in each phase of the PBL process is shown to be strongly influenced by the earlier phase, thus providing support for the PBL cycle of initial problem analysis, followed by self-directed learning, and a subsequent reporting phase as described by various authors. Alternative hypotheses where students' achievement is predicted only by collaborative learning or self-directed learning were shown to be insufficient to explain the data observed. Secondly, we have identified two distinct phases of initial terminology articulation and a later terminology repetition in the PBL process, thus providing further insight into the process of learning in PBL through a semi-naturalistic approach, instead of depending on student self-report. Lastly, we have described a useful and efficient method to keep track of students' learning throughout the PBL process.

**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

Molecular Cell Biology problem that students worked on for the day.

Made for the job

Living things use the DNA molecule to store their genetic information and to pass this information to their offspring.

Analyze the structure of DNA, and determine why it is suitable to assume this role.

## References

- Arbuckle, J. L. (2006). *Amos 7.0 user's guide*. Chicago: SPSS.
- Barrows, H. S. (1986). A taxonomy of problem-based learning methods. *Medical Education*, 20, 481–486.
- Barrows, H. S. (1988). *The tutorial process*. Springfield, Illinois: Southern Illinois University School of Medicine.
- Barrows, H. S. (1996). Problem-based learning in medicine and beyond: A brief overview. In L. Wilkerson & W. H. Gijsselaers (Eds.), *New directions for teaching and learning (Vol. 68)* (pp. 3–11). San Francisco: Jossey-Bass Publishers.
- Bonwell, C. C., & Eison, J. A. (1991). *Creating excitement in the classroom*. Washington, DC: George Washington University, School of Education and Human Development.
- Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behaviour*, 11, 717–726.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.
- Cobb, P. (1994). Where is the mind? Constructivist and sociocultural perspectives on mathematical development. *Educational Researcher*, 23, 13–20.
- Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 8, 240–247.
- De Grave, W. S., Boshuizen, H. P. A., & Schmidt, H. G. (1996). Problem based learning: Cognitive and metacognitive processes during problem analysis. *Instructional Science*, 24(5), 321–341.
- De Grave, W. S., Schmidt, H. G., & Boshuizen, H. P. A. (2001). Effects of problem-based discussion on studying a subsequent text: A randomized trial among first year medical students. *Instructional Science*, 29(1), 33–44.
- Dolmans, D., & Schmidt, H. G. (2006). What do we know about cognitive and motivational effects of small group tutorials in problem-based learning? *Advances in Health Sciences Education*, 11(4), 321–336.

- Driver, R., Asoko, H., Leach, J., Mortimer, E., & Scott, P. (1994). Constructing scientific knowledge in the classroom. *Educational Researcher*, 23, 5–12.
- Gallagher, S. A., Stepien, W. J., & Rosenthal, H. (1992). The effects of problem-based learning on problem solving. *Gifted Child Quarterly*, 36, 195–200.
- Geerligs, T. (1995). Students' thoughts during problem-based small-group discussions. *Instructional Science*, 22, 269–278.
- Gijselaers, W. H., & Schmidt, H. G. (1990). Towards a causal model of student learning within the context of a problem-based curriculum. In Z. Norman, H. G. Schmidt, & E. Ezzat (Eds.), *Innovation in medical education. An evaluation of its present status* (pp. 95–113). New York: Springer Publishing Company.
- Glaser, R., & Bassok, M. (1989). Learning theory and the study of instruction. *Annual Review of Psychology*, 40, 631–666.
- Hak, T., & Maguire, P. (2000). Group process: The black box of studies on problem-based learning. *Academic Medicine*, 75(7), 769–772.
- Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review*, 16(3), 235–266.
- Hmelo-Silver, C. E., & Barrows, H. S. (2006). Goals and strategies of a problem-based learning facilitator. *The Interdisciplinary Journal of Problem-based Learning*, 1(1), 21–39.
- Hmelo-Silver, C. E., & Barrows, H. S. (2008). Facilitating collaborative knowledge building. *Cognition and Instruction*, 26(1), 48–94.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55.
- Kintsch, W., & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85(5), 363–394.
- Kolodner, J. L., Camp, P. J., Crismond, D., Fasse, B., Gray, J., Holbrook, J., et al. (2003). Problem-based learning meets case-based reasoning in the middle-school science classroom: Putting learning by Design™ into practice. *Journal of the Learning Sciences*, 12(4), 495–547.
- Meyer, B. J. F. (1985). Prose analysis: Purposes, procedures, and problems. In B. K. Britton & J. Black (Eds.), *Analyzing and understanding expository text* (pp. 11–64, 269–304). Hillsdale, N.J.: Erlbaum.
- Michael, J. (2006). Where's the evidence that active learning works? *Advances in Physiology Education*, 30, 159–167.
- Palincsar, A. S. (1998). Social constructivist perspectives on teaching and learning. *Annual Review of Psychology*, 49, 345–375.
- Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education*, 93(3), 223–231.
- Rumelhart, D. E., & Norman, D. A. (1978). Accretion, tuning, and restructuring: Three modes of learning. In J. W. Cotton & R. L. Klatzky (Eds.), *Semantic factors in cognition* (pp. 37–54). Hillsdale, NJ: Erlbaum.
- Schiefele, U., & Krapp, A. (1996). Topic interest and free recall of expository text. *Learning and Individual Differences*, 8(2), 141–160.
- Schmidt, H. G. (1983). Problem-based learning: Rationale and description. *Medical Education*, 17, 11–16.
- 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-Silver (Eds.), *Problem-based learning: A research perspective on learning interactions* (pp. 19–52). Mahwah, NJ: Lawrence Erlbaum.
- Schmidt, H. G., De Volder, M. L., De Grave, W. S., Moust, J. H. C., & Patel, V. L. (1989). Explanatory models in the processing of science text: The role of prior knowledge activation through small-group discussion. *Journal of Educational Psychology*, 81(4), 610–619.
- Schmidt, H. G., Van der Molen, H. T., Te Winkel, W. W. R., & Wijnen, W. H. F. W. (2009). Constructivist, problem-based, learning does work: A meta-analysis of curricular comparisons involving a single medical school. *Educational Psychologist*, 44(4), 227–249.
- Van den Hurk, M. M., Dolmans, D., Wolfhagen, I., & Van der Vleuten, C. P. M. (2001). Testing a causal model for learning in a problem-based curriculum. *Advances in Health Sciences Education*, 6(2), 141–149.
- Visschers-Pleijers, A. J., Dolmans, D., Wolfhagen, I. H., & Van der Vleuten, C. P. (2004). Exploration of a method to analyze group interactions in problem-based learning. *Medical Teacher*, 26(5), 471–478.
- Visschers-Pleijers, A. J., Dolmans, D., de Leng, B. A., Wolfhagen, I., & Van der Vleuten, C. P. M. (2006). Analysis of verbal interactions in tutorial groups: A process study. *Medical Education*, 40(2), 129–137.
- Yew, E. H. J., & Schmidt, H. G. (2008). Evidence for constructive, self-regulatory, and collaborative processes in problem-based learning. *Advances in Health Sciences Education*, 14(2), 251–273.

- Yew, E. H. J., & Schmidt, H. G. (2010). What students learn in problem-based learning: A process analysis. Manuscript submitted for publication.
- Zimmerman, B. J. (1990). Self-regulated learning and academic-achievement—an overview. *Educational Psychologist*, *25*(1), 3–17.
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, *41*(2), 64–70.