# Effects of Problem-Based Learning when Taking into Account Time and Type of Assessment

This chapter is in preparation as:

*Wijnen, M., Loyens, S. M. M., Smeets, G., Kroeze, M. J., & Van der Molen, H.T. Effects of Problem-based Learning when Taking into account Time and Type of Assesment. Manuscript in preparation.*

## ABSTRACT

The present study aims to shed light on the question whether PBL is effective in terms of knowledge acquisition, when taking into account both time and type of assessment. In a randomized, controlled experiment, participants were assigned to either a PBL or a lecture condition. They learned a topic about Dutch criminal law and were tested on factual knowledge, application of knowledge, and transfer of knowledge, in both an immediate and delayed (i.e., one week later) test. Results showed no effects on knowledge retention over time, possibly due to the short time between immediate and delayed tests. Further, participants in the lecture condition outperformed PBL participants on factual knowledge questions, although performance in both conditions were sufficient. Participants in the PBL condition, however, performed better on application of knowledge assignment. Directly transmitting information to students appears helpful when students need to acquire basic knowledge, but when they need to relate their knowledge to a certain situation, the processes in PBL, e.g., activation of prior knowledge and elaboration, seem to be crucial. No differences regarding transfer of knowledge were found. The findings are both of theoretical and practical value.

## INTRODUCTION

In academic programs lectures are often a core instructional method. Providing such lectures is considered a very efficient way of teaching since a lot of information can be transmitted from the teacher to a large group of students at once. However, there are several limitations bound to lectures as well. For example, students often fail to reach higher order thinking skills (e.g., application of knowledge) due to the passive nature of lectures (Bligh, 2000; White et al., 2016). These skills are, however, very important in higher education and in life after university. Educational methods in which students are required to be *actively* involved in learning, like student-centered methods, aim to foster these skills. Problem-based learning (PBL) is an example of such an approach. The present study focuses on the differences between instruction by PBL and by lectures on student performance. First, we will go into depth about the origin and process of PBL.

### Problem-Based Learning

At the McMaster University Medical School in Canada in the 1960s, students experienced difficulties with understanding complex topics, were less motivated, and did not see the relevance for their profession (Barrows, 1996). To motivate and help students, working with and discussing realistic problems in small groups was introduced as instructional method. This was referred to as PBL. Since its origin (Norman & Schmidt, 1992; Schmidt, 1983), PBL has been implemented all over the world and over the past decades, different variations of PBL have evolved (Barrows, 1996; Loyens, Paas, & Kirschner, 2012). Despite various types, the following characteristics of PBL are defined. (1) Learning is student-centered, which holds that the students themselves should take responsibility for their learning process. (2) Learning takes place in small groups. (3) The teacher acts as facilitator, meaning that he/she asks those questions that make students elaborate on information instead of providing factual information. (4) Problems that challenge and motivate students are used in the instruction, and (5) these problems should foster the development of problem-solving skills. (6) Self-directed learning should be present (Barrows, 1996).

Schmidt, Van der Molen, Te Winkel, and Wijnen (2009b) describe three perspectives of PBL. One of these perspectives considers PBL as a cognitive constructivist approach with the primary goal to build flexible mental models in learners. One of the goals of PBL, construction of an extensive knowledge structure, is in line with this perspective. The process of PBL contributes to the attempt to achieve this goal. In general, three phases are distinguished. In the initial discussion phase, students receive a problem, which is usually a description of a specific situation. Collaboratively, the problem is discussed and students try to explain it. This way, prior knowledge about the topic of the problem at hand is activated. As the problem is the starting point, students end up with questions about unclarified aspects and they formulate so called learning issues. In the second

phase, the self-study phase, students individually select and study relevant literature sources, attempting to answer the learning issues for themselves. This is the preparation for the final phase, the reporting phase. During this phase students discuss their findings together and collaboratively address the learning issues. The tutor, who is present during the initial and reporting phase, can ask in-depth questions to make students elaborate more on the material (Loyens et al., 2012; Schmidt, 1983).

In order to realize knowledge acquisition in students, the processes of activation of prior knowledge, elaboration of knowledge, and learning in a realistic context (Norman & Schmidt, 1992; Schmidt, 1983) play an important role. Students activate their prior knowledge in the initial phase when discussing the problem as starting point. When students acquire new knowledge during the self-study and reporting phase, it is easier for them to connect new knowledge to existing knowledge in memory. This is also called the process of elaboration and takes place trough discussion, fostering knowledge retrieval (Norman & Schmidt, 1992; Schmidt, 1983). Hence, this assumes that students in PBL acquire more knowledge of the PBL instruction than students in more traditional, lecture-based environments.

### Knowledge Acquisition and Retention

Several studies have focused on the effectiveness of PBL with regards to knowledge acquisition. In the majority of these studies, PBL students were compared with students from more conventional, lecture-based educational methods on their academic performance, e.g. course exams. Several meta-analyses that contain these effect studies exist, however with inconclusive results. Most meta-analyses demonstrated in general no differences between PBL and non-PBL students, or even negative effects of PBL on their immediate knowledge acquisition (Albanese & Mitchell, 1993; Dochy, Segers, Van de Bossche, & Gijbels, 2003; Schmidt, Van der Molen, Te Winkel, & Wijnen, 2009b; Vernon & Blake, 1993). However, a recent meta-analysis of Dağyar and Demirel (2015) demonstrated that PBL students obtain better academic achievements than students of conventional curricula.

The meta-analysis of Dochy et al., (2003) indicated the importance to also explore effects of PBL over time. PBL does seem to have a positive effect with regards to knowledge *retention* on the long-term: PBL students perform better on delayed tests and hence retain more knowledge over time compared to their non-PBL counterparts (Dochy et al., 2003; Strobel & Barneveld, 2009). In short, studies demonstrate that PBL is not necessarily beneficial in terms of immediate knowledge acquisition, but retention on the long-term appears to be better among PBL students.

### Application and Transfer of Knowledge in Problem-Based Learning

Besides the importance of retention period in assessment, the type of assessment should be taken into account as well. The meta-analysis of Gijbels, Dochy, Van den Bossche, and

Segers (2005) showed that PBL students, compared to students of traditional curricula, perform better when assessments focus on higher levels of knowledge structures. These levels contain the understanding of principles that link concepts and the application of knowledge (i.e., procedure; Sugrue, 1993). The lower level of knowledge structure on the other hand means understanding of concepts (e.g., factual knowledge). Similar results were found in an experiment by Masek and Yamin (2012): students taught by lectures performed better on the understanding of concepts, while PBL students acquired more knowledge regarding principles and procedures (Masek & Yamin, 2012). An explanation for this is that the instructions of PBL are more in line with the higher level of knowledge structures. For example, in PBL, learning takes place in a realistic context (i.e., the problem presented to the students) that requires students to link the course material to real-life situations and therefore apply their knowledge to a certain extent. Further, elaborating and discussing the material might contribute to application of knowledge as well, because students refer back to the problem they started with. These aspects might even help students to *transfer* the knowledge to new situations.

Transfer is the process in which students apply the knowledge they have learned in a different and novel context (Perkins & Salamon, 1992). Transfer is a very important aspect of education (Perkins & Salamon, 1992), as students need to be able to apply the knowledge they have learned in real-life situations, as well as in their future profession (Pugh & Bergin, 2006). Despite its importance, transfer is a difficult process that does not happen automatically (Norman, 2009). In order for transfer to take place, students need to be able to recognize and understand the underlying principles in different situations or contexts. However, not all contexts and situations look similar and it is difficult for students, especially for novices, to recognize the deeper, underlying principles (Norman, 2009).

It could be argued that PBL fosters transfer, because students start their learning process in a realistic context. The learning material is integrated with a realistic and complex problem, making it easier to relate acquired knowledge to new situations. Previous studies that have focused on PBL's effect on transfer (Bergstrom, Pugh, Philips, & Machlev, 2016; Pease & Kuhn, 2011; Wirkala & Kuhn, 2011) found that indeed students in PBL are better in applying their knowledge in new and different situations and hence, that PBL seems to stimulate transfer.

### The Present Study

As reported above, a lot of research has been done on the effectiveness of PBL on knowledge acquisition. However, one general shortcoming of the existing studies that the present study tries to overcome is the lack of controlled experiments (Kirschner, Sweller, & Clark, 2006). The majority of the PBL effect studies are conducted in existing curricula and courses. Although this is highly ecologically valid, there are many external factors that might influence the results such as the group composition of the tutorial groups

and the tutor. Wijnen, Schaap, and Loyens (2016) aimed to overcome this by conducting a randomized controlled lab-experiment on the effectiveness of PBL. Participants were taught on a psychology topic either by PBL, a lecture, or through self-study. They were tested on their acquired knowledge on a multiple-choice (MC), factual knowledge test, both immediate as delayed. Results showed that participants in the PBL condition outperformed those in the lecture condition on knowledge acquisition. However, besides the issue of controlled experiments, two remarks about PBL effect studies remain.

First, the majority of the PBL effect studies are conducted in the area of medical education or social sciences (e.g., psychology). However, in order to generalize findings, studies should focus on different disciplines as well. The present study tries to overcome this by focusing on a different area, legal education. To our knowledge, effects of PBL have not often been investigated within this discipline. The second remark relates to the type of assessments used in PBL effect studies. A limitation of past studies is that the different levels of knowledge structures have seldom been measured at the same time. The present study will therefore focus on factual knowledge, application of knowledge, and transfer of knowledge and hence take several levels (i.e., lower and higher levels) into account.

In the present study a controlled experiment was conducted. Participants were randomly assigned to either a PBL or lecture condition and were compared on the performances on different types of assessment mentioned above that were tested on both an immediate as delayed test. Figure 6.1 depicts the procedure of the experiment. In the Method section the procedure is discussed in more depth. The first research question was "What is the influence of PBL on a) knowledge acquisition and b) knowledge retention?" The influence of PBL has been studied on both an immediate and delayed test to measure respectively knowledge acquisition and retention. Previous studies showed mixed findings regarding PBL's effect on knowledge acquisition (Albanese & Mitchell, 1993; Dağyar & Demirel, 2015; Dochy et al., 2003; Schmidt et al., 2009b; Vernon & Blake, 1993). Therefore, for the first part of the first research question, no specific hypothesis is formulated, but we kept the question more explorative. Moreover, acquisition of knowledge is expected to differ between types of assessment, which will be outlined in the next paragraph. Regarding the second part of the first research question (hypothesis 1b), it is hypothesized that PBL students outperform those in the lecture condition on the delayed test despite the type of assessment (Dochy et al., 2003; Strobel & Barneveld, 2009).

The second research question focused on the type of assessment: "What is the influence of PBL on a) factual knowledge, b) application of knowledge, and c) transfer of knowledge?" It was hypothesized that regarding the factual knowledge questions (hypothesis 2a), participants in the lecture condition performed better (Masek & Yamin, 2012). However, we expected that PBL students performed better on the application

assignment (hypothesis 2b; Gijbels et al., 2005; Masek & Yamin, 2012) and the transfer assignments (hypothesis 2c; Bergstrom et al., 2016; Pease & Kuhn, 2010; Wirkala & Kuhn, 2011).

## METHOD

### Participants

Participants were undergraduate psychology students from a Dutch university and so they were novices in the area of Dutch criminal law and were expected to have a similar level of prior knowledge. In total, 67 students participated (17.9% male). Age ranged from 18 to 25 and the mean age was 20.0 years ($SD$ = 1.56). Participants were randomly assigned to either the lecture condition ($n$ = 33; 21.2% male) or PBL condition ($n$ = 34; 14.7% male), mean age = 19.9 ($SD$ = 1.51) and mean age = 20.0 year ($SD$ = 1.62) respectively. There were no significant differences between both conditions regarding age ($t(65)$= -.70, $p$ = .488) and gender ($\chi^2(1)$ = .48, $p$ = .487). In return for participation, students earned credit points.

### Material

Participants learned about Dutch criminal law topics "self-defense" and "unreasonable use of self-defense". This topic is part of the academic Dutch law program at the university under study. It is briefly brought up in the first academic year during the introductory course of criminal law, and more in-depth during the second academic year, in the follow-up course of Dutch criminal law. Since participants were psychology students, there were no conflicts between the content of the topic to-be-learned in the experiment and the content of the curriculum of the participants. Several materials were deployed in the learning phases of both conditions (i.e., the problem for the PBL condition, the text, and the test). Two independent experts on Dutch criminal law were involved in the development of some of the materials and conducting the experiment. The first expert helped with the construction of the lecture for the lecture condition and the test to measure the effects of the intervention. Additionally, this expert assisted with rating the answers to the open questions. The second expert acted as lecturer in the lecture condition and as tutor in the PBL condition. By doing so, we controlled for possible differences in instructors. The expert was briefed in advance and the first author was present during the experiment in both conditions. Both experts involved were young jurists, graduated in Dutch criminal law and were employed at the university under study for several years.

***Lecture***

Participants in the lecture condition received a lecture of 45 minutes. The lecture contained 26 Powerpoint slides. The three main topics that were addressed were self-defense, unreasonable use of self-defense, and *culpa in causa* (i.e., guilt by cause). The lecture started with defining self-defense according to Dutch law. Next, all required conditions to appeal to self-defense were mentioned along with examples (e.g., the attack needs to be *immediate* in order to appeal to self-defense). After self-defense, unreasonable use of self-defense was addressed: how it is defined in Dutch law, the conditions that are required, and again some examples. Next, an explanation was given of *culpa in causa* (i.e., appeal to self-defense and unreasonable use of self-defense will not succeed when someone is seeking the confrontation). The final part of the lecture focused on a (fictive) news article (*"Failed drug deal"*) relating to a situation in which someone defended himself and appealed for self-defense. This news article was exactly the same as the problem in the PBL condition. This news article is described in more detail below (i.e., problem). The conditions for self-defense in this particular case were explained to the participants.

The lecturer received the instruction to give the lecture in a realistic way, to resemble the existing educational practice as closely as possible. The instructor asked several questions during the lecture to the participants (e.g., "Can you try to explain in your own words what is said here on the slide?", "Can you give an example of an attack that is out of proportion?"). When discussing the news article, the lecturer asked participants the question whether they thought that the person in the article could appeal to (unreasonable use of) self-defense. Some of the participants were asked to explain themselves and some discussion arose in the audience.

***Problem***

The problem in the PBL condition was a fictive news article titled *"Failed drug deal"*. This news article described a situation in which a drug dealer was robbed of his drugs and drugs money by another man. The brother of the drug dealer wanted revenge for his brother and – with a gun – he left his house to search for the man who robbed his brother. When they met each other, the man who stole the drugs and money was running up to the brother with a knife in his hand. The brother was not able to run away and he shot the other man in the chest, with immediate death as a consequence. The lawyer of the drug dealer's brother claimed that shooting the man was self-defense. In the initial phase, participants needed to discuss whether they agree with the lawyer, and whether they thought an appeal to self-defense will succeed. The Seven-Jump method (Schmidt, 1983) was used to shape the PBL process.

This problem is part of the curriculum of Dutch law at the Erasmus School of Law. In the criminal law course in the second academic year, self-defense and unreasonable

use of self-defense is part of the course. All designed problems of all courses at the law program under study are checked and provided with feedback by PBL experts. In addition, before these problems are used in the law program, they are tested out with a small sample of students to check whether they work sufficiently. If that is not the case, adjustments are made. Therefore we assume that the quality of the problem at hand was good.

***Text for Self-Study***

In both conditions, there was a period of 45 minutes of self-study in which participants had the opportunity to study. A text was provided in which self-defense, unreasonable use of self-defense, and *culpa in causa* were explained, along with all of its conditions and requirements. This text was copied from a study book that is used in the Dutch law curriculum of the university under study. The text contained 13 pages and was written in Dutch. The same topics that were addressed in the lecture condition were cited in the text. To be more specific, the lecture was based on the text. Besides the study text, participants received a copy of the relevant Dutch law articles (i.e., in Dutch: Art. 41 lid 1 Sr, and Art. 42 lid 2 Sr).

***Test***

The test used in the experiment contained three parts. First, 10 MC questions that measured factual knowledge. Second, an assignment to measure application of knowledge, and third an assignment that measured transfer of knowledge. The test was the same in both conditions and both test phases. The first author in collaboration with one of the experts developed the test. The test was based on the text and the questions had 3 answer options each. An example of a MC question was: "Immediate attack is a requirement of: A) Self-defense, B) Unreasonable use of self-defense, C) Both self-defense as unreasonable use of self-defense. (C is the correct alternative)" Participants could obtain one point for each correctly answered MC question. This resulted in a minimum obtainable score of 0 and a maximum obtainable score of 10 for this part of the test.

The two assignments that were administered, one for application of knowledge and transfer of knowledge, both contained a news article in which a situation was described. The application assignment represented the news article that all participants saw and discussed before (i.e., as problem in the PBL condition and final slide in lecture condition). The question belonging to the article was "Motivate whether you think an appeal to self-defense / unreasonable use of self-defense will succeed in this case? Explain, step by step, how you came to your conclusion." A total score of 10 points could be obtained for this assignment.

The second assignment intended to measure the transfer of knowledge. A new fictive news article was presented here with the following situation: A woman and her three-

year old daughter were attacked by the husband of the woman (and father of the child). The woman and husband started a discussion, but this resulted in physical assault by the husband (i.e., grabbing the woman and shaking her). Moreover, when the daughter started to cry, the husband roughly grabbed the child's arm. The woman tried to pull him away from their daughter and he furiously ran up towards her. The woman grabbed a rolling pin lying on the kitchen counter next to her and she hit her husband on the head. He fell to the ground and did not move anymore. After this, she hits him again with the rolling pin. The question belonging to this news article was "Motivate whether you think an appeal to self-defense / unreasonable use of self-defense will succeed in this case? Explain, step by step, how you came to your conclusion." A total score of 10 points could be obtained for this assignment as well.

A model answer for both assignments was developed by the first author and expert. In this answer, the correct steps for coming to the right conclusion were mentioned with the number of points earned for each part. Answers to the assignments were rated based on these model answers, by the first expert and partly by the first author. Both raters were blind for participants and the condition of the participants. Interrater reliability turned out to be very high, as an intraclass correlation coefficient (ICC) of .95 was reached.

**Design of the Experiment and Procedure**

The experiment consisted of a learning phase, in which instructions differed between conditions (i.e., lecture vs. PBL), and two test phases (i.e., immediate test and delayed test) that were the same in both conditions. No pre-test was administered, in order to prevent priming of knowledge.

In the lecture condition participants started the experiment with a lecture of 45 minutes. During the lecture, participants were allowed to ask questions and to take notes. After that, a 45 minute period of self-study started, in which participants had the opportunity to study the text. Again, they were allowed to take notes. The first, immediate test phase started right after self-study. The experimenter collected the texts and notes of all participants before administrating the test. Participants had a total of 30 minutes to fill out the test. After a week, participants returned for the second, delayed test phase in which they filled out the same test for which they had 30 minutes. In order to test the retention of acquired knowledge after one week participants were not informed in advance that the test would be administered again the second time they returned for the experiment.

In the PBL condition, participants were assigned to one of a total of four PBL groups. Each group consisted of about nine participants and one tutor. The tutor could ask in-depth questions about the problem, making students elaborate more on the material. Moreover, the tutor monitored whether everyone participated actively in the discus-

sions. Each PBL group started the experiment with the initial phase, which took about 15 minutes. During the initial phase, participants discussed the problem *"Failed drug deal"*. The PBL process in the experiment followed the Seven-Jump method (Schmidt, 1983) and hence the first five steps were applied in the initial phase (i.e., clarifying terms and concepts, defining the problem, brainstorming, problem analyzing, and formulating learning issues). The formulated learning issues were comparable in all groups and they came down to: "What is self-defense?", "What if self-defense is out of proportion?" and "What happens when one appeals to self-defense, but he/she actually seeks confrontation?" These learning issues referred to the three main topics in the text. After the initial phase, a 45 minute period of self-study started, which is the sixth step in the Seven-Jump method. Participants were allowed to take notes. The reporting phase started afterwards (i.e., seventh step) and took 30 minutes. One of the participants volunteered or was chosen to guide the discussion in the reporting phase. Participants answered the learning issues during the reporting phase and the instructor was told to make sure that in all PBL groups, a connection should be made to the problem. After the reporting phase, the experimenter collected all texts and written notes before administering the test. Similar as in the lecture condition, participants had 30 minutes to fill out the test. One week later, participants returned to fill out the same test again (30 minutes) in the second test phase. Participants were not informed in advance that they would be asked to do the test again.

In both conditions, the first author was present during the whole experiment. In the lecture condition, the lecturer left after giving the lecture, and in the PBL condition, the tutor was only present during the initial and reporting phase. In addition, the instructor was unaware of the exact content of the given test. Time on task was equal in both conditions: the total time of the experiment was two and a half hours. The procedure is depicted in Figure 6.1.

### Data analysis

Three Mixed Analysis of Variance (ANOVA's) were conducted, one for each type of assessment. The first regarding scores on the factual, MC questions, the second for the application assignment, and the third for the transfer assignment. In all Mixed ANOVA's, the between-subjects factor was condition (lecture vs. PBL) and the within-subjects factor was time (immediate vs. delayed). Knowledge *acquisition* for each type of assessment was operationalized as main effect of instruction type. Knowledge *retention* was operationalized by the interaction of time and instruction type. This interaction gave evidence of the knowledge that participants retained from the immediate to the delayed test and whether that differed between conditions. Effects were reported as significant when $p < .05$. Partial eta-squared effect sizes indicated the size of the effect. A value of partial $\eta^2 > .01$ is considered small, a partial $\eta^2 > .06$ is considered medium and a partial $\eta^2 > .14$ is considered large.

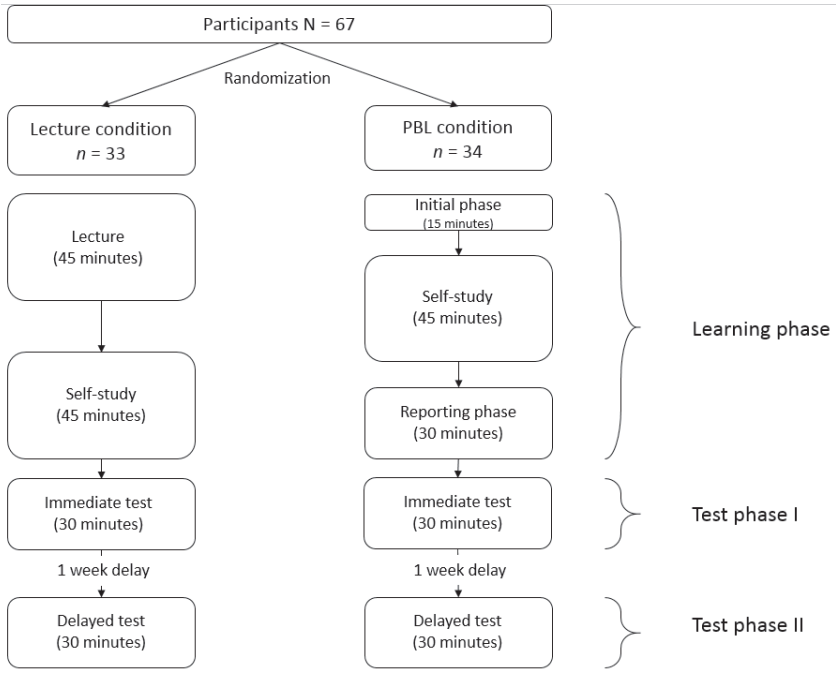


**Figure 6.1.** Experimental procedure

## RESULTS

Before conducting the analyses, assumptions for mixed ANOVA's were checked and met (i.e., normality, Leven's tests). However, the reliability of the MC test appeared very low (Cronbach's alpha of .13). Therefore, results on the first mixed ANOVA should be interpreted with caution. Two participants did not show up on the delayed test phase and they were left out of further analyses, resulting in a total number of 65 participants. Table 6.1 provides the mean scores in both conditions on all three assessment types. Mean scores demonstrated that performance on the MC questions was sufficient in both conditions (a mean score of 6-7 out of 10), however, scores on the application and transfer assignments were rather low (a mean score of 2-3 out of 10).

Results are discussed for each assessment type separate. First, the effect of time is mentioned, followed by the effect of instruction type, and ending with the interaction effect.

**Table 6.1.** Mean scores on the immediate and delayed tests for both conditions

| | | Immediate test | Delayed test |
|---|---|---|---|
| MC questions | PBL | 6.53 (*SD* = 1.30) | 6.19 (*SD* = 1.62) |
| | Lecture | 7.18 (*SD* = 1.19) | 6.85 (*SD* = 1.37) |
| Application assignment | PBL | 3.08 (*SD* =1.61) | 2.88 (*SD* = 1.08) |
| | Lecture | 2.33 (*SD* = .97) | 2.27 (*SD* = .83) |
| Transfer assignment | PBL | 2.81 (*SD* = 1.63) | 2.48 (*SD* = 1.44) |
| | Lecture | 3.30 (*SD* = 1.57) | 3.17 (*SD* = 1.41) |

*Note*. Scores could range from 0 to 10 on all types of assessment.

### Factual knowledge MC questions

Although both for the PBL and the lecture condition the mean scores were lower at the delayed test than at the immediate test, there was no main effect of time, $F(1, 63) = 2.54$, $p = .116$, partial $\eta^2 = .04$. A significant main effect of instruction type showed up, in favor of the participants in the lecture condition, $F(1, 63) = 6.04$, $p = .017$, partial $\eta^2 = .09$ (medium effect), which supports hypothesis 2a. Furthermore, there was no interaction effect, $F(1, 63) = .00$, $p = .981$, partial $\eta^2 = .00$, so performance over time was similar in both conditions. Hence, regarding factual knowledge questions, hypothesis 1b on knowledge retention was not confirmed. Still, as mentioned before, results on the MC questions should be interpreted with caution due to the low reliability.

### Application of knowledge assignment

No main effect of time was found for the open-ended question $F(1, 63) = .62$, $p = .431$, partial $\eta^2 = .01$. There was a significant main effect of instruction type, $F(1, 63) = 8.35$, $p = .005$, partial $\eta^2 = .18$ (large effect). Giving support to hypothesis 2b, participants in the PBL condition outperformed participants in the lecture condition. Further, no interaction effect was present, $F(1, 63) = .18$, $p = .670$, partial $\eta^2 = .00$. Hypothesis 1b was not confirmed with regards to the application of knowledge, as PBL students did not retain more knowledge over time.

### Transfer of knowledge assignment

There was no main effect of time for the transfer question, $F(1, 63) = 1.36$, $p = .249$, partial $\eta^2 = .02$. Also, no effect of instruction type appeared, $F(1, 63) = 3.39$, $p = .071$, partial $\eta^2 = .05$, which was contrary to hypothesis 2c that stated that participants in the PBL condition would outperform those in the lecture condition. Finally, no interaction effect was found, $F(1, 63) = .23$, $p = .632$, partial $\eta^2 = .00$. Again, hypothesis 1b was not confirmed with regards to the transfer of knowledge.

## DISCUSSION

The present study focused on the effectiveness of PBL on time and type of assessment. Knowledge acquisition and knowledge retention were studied on three assessment types: factual knowledge MC questions, one assignment intending to measure the application of knowledge, and a second assignment intending to measure transfer of knowledge. A controlled experiment was conducted, in which participants were randomly assigned to either a PBL or a lecture condition and learned about a Dutch law topic. Measurements took place immediately after the instruction and one week after that.

### Time of Assessment: Knowledge Acquisition and Retention

#### *Knowledge acquisition*

The first research question focused on the influence of PBL on immediate knowledge acquisition and knowledge retention. No specific hypothesis with regards to knowledge acquisition was formulated because of inconclusive findings in existing literature (Albanese & Mitchell, 1993; Dağyar & Demirel, 2015; Dochy et al., 2003; Schmidt et al., 2009b; Vernon & Blake, 1993). Moreover, it was expected that knowledge acquisition would differ between the PBL and lecture condition for the types of assessment (Gijbels et al., 2005). The latter was indeed shown in the results. Participants in the lecture condition outperformed those in the PBL condition on the factual knowledge questions, while it was the other way around for the application of knowledge assignment. These results are discussed below.

#### *Knowledge retention*

Further, it was hypothesized that retention of knowledge over time would be higher when participants received instructions through PBL, on all types of assessment. In PBL there is emphasis on elaboration that is assumed to help students remember more of the learned knowledge (Dochy et al., 2003; Schmidt, 1983; Strobel & Barneveld, 2009). In addition, it was expected that participants in the lecture condition would forget more of the acquired knowledge. Results were however contrary to the hypothesis, as no interaction effects were shown for any of the assessment types. Knowledge loss over time was similar for all participants, irrespective of the instructional method used.

A possible explanation for these findings is that the time between the immediate and delayed test was only one week. We can relate this to the finding that for none of the assessment types there was a main effect of time, meaning that performance stayed equal over time. It could be argued that a week is too short to detect the effect of PBL on the long-term. The study of Capon and Kuhn (2004), for example, did show an effect

of PBL over time, but their test was administered six weeks after the learning phase. However, since the time on task in this experiment was shorter than a regular lecture or PBL tutorial meeting, the time between tests was adapted to this and kept on a week.

### Type of Assessment: Factual Knowledge, Application of Knowledge and Transfer of Knowledge

The second research question focused on the level of knowledge that was assessed. We discriminated between factual knowledge, application of knowledge, and transfer of knowledge. It was expected that differences showed up for type of assessment. Our specific hypotheses stated that participants in the lecture condition outperformed participants in the PBL condition on factual knowledge (hypothesis 2a), however, that it was expected that this was the other way around for application of knowledge (hypothesis 2b), and transfer of knowledge (hypothesis 2c). Results met some of our expectations.

#### *Factual knowledge*

It was found that participants in the lecture condition outperformed participants in the PBL condition on factual knowledge, measured by 10 MC questions, confirming hypothesis 2a. This result is in line with findings of for example Capon and Kuhn (2004) in which students in a lecture-based environment performed better when assessment focused on the understanding of concepts (i.e., basic knowledge). Apparently, transmitting information directly from a teacher during a lecture is beneficial for the understanding and reproduction of basic knowledge. Although the mean scores are higher in the lecture condition than in the PBL condition, if the caesura between an insufficient and a sufficient score is put at 6.0, performances on the MC questions in general were sufficient in both conditions. Again, the reliability of the MC test turned out very low. Results on this part of the test should therefore be interpreted with caution.

#### *Application of knowledge*

Regarding the assignment that required application of the learned knowledge, participants in the PBL condition scored higher than participants in the lecture condition, supporting our hypothesis. This result is in line with findings of Masek and Yamin (2012), in which PBL students were better in applying their knowledge than students taught by lectures. In the present study, the context of the application assignment was made familiar for all participants during the learning phase. There was equal exposure to the context of the assignment in both conditions (as problem in the PBL condition and as news article on the final slide of the lecture in the lecture condition). Despite the fact that in the lecture condition participants received an explanation of the situation described, it did not help them to reach the same level of performance as participants in the PBL condition. This indicates that the specific processes that occur in PBL, such as

activation of prior knowledge, elaboration, and collaborative discussion, contribute to the application of knowledge (Capon & Kuhn, 2004; Schmidt, 1983). During PBL discussions, students explain and elaborate on the literature collaboratively and they refer to the problem in the reporting phase. Hence, students integrate the described situation of the problem at hand with the knowledge they have acquired. This will help application of knowledge at a later point in time.

It should be noted that in general performance on the application assignment was quite low (a score of about 3 out of 10). Applying knowledge appeared a difficult process for participants. This might be due to the limited amount of time of the experiment. The total time of the experiment took 2.5 hours, which is perhaps too short for novice students to learn sufficiently about a topic in a new area.

***Transfer of knowledge***

With regards to performance on transfer of knowledge, operationalized by the last assignment, there were no differences between participants in the PBL and lecture condition. This was not in line with our hypothesis, as it was expected that instructions by PBL would result in better transfer of knowledge. It was expected that the processes mentioned before – activation of prior knowledge and elaboration – would be beneficial for transfer as well. Prior studies on PBL's effects on transfer tasks demonstrated higher performances of PBL students compared to students in traditional learning environments (Bergstrom et al., 2016; Pease & Kuhn, 2010; Wirkala & Kuhn, 2011).

There is a possibility that the transfer assignment was too challenging for participants. To transfer knowledge to new situations, although very important, is in general a difficult process for students to master (Norman, 2009; Perkins & Salamon, 1992). For example, in his article, Norman (2009) described that only about 10-30% of the medical students is able to do this. This can be observed in the low mean scores on the transfer assignment (score around 3 out of 10). The short time span of the experiment might also provide an explanation here. Two and a half hours is probably too short for novice students to master the knowledge and be able to transfer this to new situations.

## LIMITATIONS

Several limitations of this study have to be mentioned. First, the total time of the experiment was relatively short compared to regular educational practices. For example, the time on the PBL process in this study is shorter than the normal PBL process at the university under study (i.e., 45 minutes for initial and reporting phase vs. two and a half hours for initial and reporting phase for one PBL session in real education). The same applies to the lecture time (i.e., 45 minutes vs. approximately two hours). A shorter time

for a lecture might be more beneficial for participants, because it is easier to keep their attention, while for the PBL session, the shorter time can be more of a disadvantage. A second limitation was that in the PBL condition, participants were not able to select their own literature sources. In the experiment, one text was provided for self-study. Despite the fact that this is in contrast with a realistic PBL setting, providing one text for all participants made sure there was controlled for similar knowledge intake among participants. A third important limitation is the very low reliability of the factual knowledge MC test. The low reliability has to do with the limited number of questions that could be developed about the single subject that was taught to the participants. This makes it difficult to interpret any of the findings regarding the factual knowledge assessment.

## CONCLUSION

The present study contributes to existing literature on PBL effect studies in, because of the controlled experimental nature of this study (opposed to effect studies in existing curricula) and the discipline in which it took place (legal education opposed to medical education). The findings of this study show the importance of type of assessment used when investigating the effectiveness of PBL compared to lectures. When testing factual knowledge, lectures seem to be more effective, although the mean performance on this assessment type is sufficient for both instruction types and the reliability of this part of the test was low. However, when students need to apply their knowledge to a realistic situation, PBL instructions seem to be somewhat more beneficial. In this study we have opposed PBL to lectures. Though, as is shown by the results of this study, both methods can be used to support each other and in practice this is often the case. A combination of lecturing for acquiring basic, factual knowledge and PBL for getting deeper understanding and application of knowledge, might be an advantageous way of instructing students.