Constructivist, Problem-Based Learning Does Work: A Meta-Analysis of Curricular Comparisons Involving a Single Medical School

Henk G. Schmidt, Henk T. van der Molen, and Wilco W. R. te Winkel

Department of Psychology
Erasmus University, The Netherlands

Wynand H. F. W. Wijnen

Department of Educational Development and Research
Maastricht University, The Netherlands

Effects of problem-based learning as reported in curricular comparison studies have been shown to be inconsistent over different medical schools. Therefore, we decided to summarize effects of a single well-established problem-based curriculum rather than to add up sometimes-conflicting findings from different problem-based curricula. Effect sizes were computed for 270 comparisons. The results suggest that students and graduates from the particular curriculum perform much better in the area of interpersonal skills, and with regard to practical medical skills. In addition, they consistently rate the quality of the curriculum as higher. Moreover, fewer students drop out, and those surviving need less time to graduate. Differences with respect to medical knowledge and diagnostic reasoning were on average positive but small. These outcomes are at variance with expectations voiced in recent contributions to the literature. They demonstrate that constructivist curricula can have positive effects on learning even if they deemphasize direct instruction.

In 1969, a first group of 20 medical students arrived at McMaster University, Hamilton, Ontario, to enroll in a curriculum that was quite avant-garde, even by international standards. Students engaged in a process of learning and instruction that was called “problem-based learning” (PBL). They were to work on relevant biomedical or clinical problems in collaboration with peers and guided by a tutor. The number of lectures they received each week was limited to one or two. Students were supposed to learn primarily through self-directed study, guided by the problems designed by their teachers (Hamilton, 1976; Neufeld & Barrows, 1974). In addition, the curriculum emphasized the acquisition of medical, interpersonal, and other professional competencies. Would the early developers at the time have had the appropriate terminology available, they would have described their curriculum as constructivist, promoting contextual, collaborative learning, self-regulation, and student agency (Brown, Collins, & Duguid, 1989; E. Cohen, 1994; Scardamalia & Bereiter, 1991).

Now, 40 years later, there is little doubt that PBL is quite popular as a pedagogical innovation, in particular in medical education. For example, a majority of the medical schools in the United States includes small-group tutorial sessions organized around clinical problems, and 20% of these consider themselves to be problem-based outright (Association of American Medical Colleges, 2005). In addition, most Australian medical schools have adopted PBL as their instructional method (Sanson-Fisher & Lynagh, 2005), and curricula based on these ideas have also been developed in Europe and Asia (e.g., Antepohl & Herzig, 1999; Fyrenius, Silen, & Wirell, 2007; Khoo, 2003; O’Neill, Morris, & Baxter, 2000; Tiwari, Lai, So, & Yuen, 2006). Moreover, PBL has been adopted in economics and business (Gijselaers et al., 1995), engineering (Dahlgren & Dahlgren, 2002), psychology (Reynolds, 1997), law (Moust & Nuy, 1987), and biology (Kendler & Grove, 2004). Finally, PBL
is experimented with in K-12 education in the United States (Ertmer & Simons, 2006) and The Netherlands. Hence, PBL is one of the few curriculum-wide educational innovations surviving since the 60s.

However, opinions differ as to whether PBL is an effective form of education. Some argue that, contrary to expectation, PBL has not promoted in students higher levels of knowledge (Colliver, 2000). Others maintain that PBL is a form of minimally guided instruction and is therefore less effective and less efficient than approaches that place a stronger emphasis on direct instruction of students (Kirschner, Sweller, & Clark, 2006; Klahr & Nigam, 2004).

In this contribution, we review results of a large number of studies in which the performance of students and graduates from one problem-based medical school was compared with the performance of students and graduates of various conventional medical schools. Unlike previous reviews (e.g., Albanese & Mitchell, 1993; Berkson, 1993; Colliver, 2000; Vernon & Blake, 1993), we do not confine our contribution to effects of PBL on knowledge acquisition alone. Next, studies are reviewed on the acquisition of (a) knowledge, (b) diagnostic competence, (c) interpersonal and other general professional competencies, and (d) practical medical skills. In addition, (e) student perceptions of the quality of problem-based versus conventional education are reviewed, and we pay attention to comparison studies in which not the individual students or graduates were the focus of attention but (f) the relative efficiency of the curricula as a whole, as exemplified by graduation rates and study duration of students from these curricula. The latter is particularly relevant because a recent critical analysis of PBL in this journal suggested that such approach must be less efficient than direct instruction (Kirschner et al., 2006). We propose that reporting on a broad range of outcome measures will possibly help in painting a more measured picture of PBL than was possible in recent publications on the value of this educational approach (Collier, 2002; Stanley, 2007).

As indicated, we report on studies in which students or graduates from one particular Dutch problem-based medical school were compared with their counterparts trained through conventional medical education. Although analyzing data involving only a single curriculum has obvious disadvantages, it also has a number of advantages over indiscriminately meta-analyzing effects of different curricula. These advantages are discussed next. However, first we provide, in some detail, a description of what happens to the learner in the problem-based approach to education and discuss three different interpretations of why these things happen to the learner in PBL.

**PROCESS OF LEARNING IN PROBLEM-BASED CURRICULA**

Authors generally agree that PBL has six defining characteristics: (a) the use of problems as the starting point for learning, (b) small-group collaboration, and (c) flexible guidance of a tutor. Because problems steer the learning in such curriculum, (d) numbers of lectures are limited. The latter is in line with the idea that (e) learning is to be student-initiated and that (f) ample time for self-study should be available (Barrows, 1985; Evensen & Hmel, 2000; Hmel-Silver, 2004; Schmidt, 1983, 1993).

In PBL, the problem comes first. A problem is usually a description of a set of phenomena or events observable in the real world that are in need of an explanation in terms of a theory, an underlying principle, process, or mechanism. The task of the students in PBL is to construct such theory through small-group discussion and through self-directed learning (Schmidt, 1983). The following example of a problem is taken from a 1st-year medical curriculum. The title is “Miraculous Rescue.” And the problem description goes like this:

For more than 15 minutes an eight-year-old boy, Maurice, has been lifelessly floating around in water colder than 60 degrees F. Fortunately, a passerby succeeds in bringing him out of the water. Mouth-to-mouth resuscitation is applied immediately. Everyone is astonished to notice that the boy is still alive. Presently, Maurice is on the intensive care ward of the local hospital and is out of danger. According to his doctor he is expected to recover completely. Explain why this is possible.

A second example is taken from an undergraduate problem-based course in psychology. Its title is “A Car Accident,” and here is the description of the problem:

Participants in a psychological experiment were shown a movie of two cars bumping into each other. Afterwards, some participants received the following question: “How fast did the cars drive when they hit each other?” Others received this question: “How fast did the car drive when they crashed into each other?” The first group estimated the speed to be 34 miles per hour; the second group thought that the speed was 41 miles per hour. One week later all participants were asked: “Did you also see broken glass?” (There was no broken glass). Fourteen percent of the first group said “yes.” Of the second group 32% said “yes.” Explain these findings.

As can be deduced from these examples, problems actualize important scientific ideas—here, physiological or psychological mechanisms—that students have to master as part of their education. In medical education the phenomena to be explained often take the form of signs or symptoms of a sick person. In psychology, it may be outcomes of an experiment such as the Loftus false memory findings or the Asch conformity results. In science education, natural phenomena observed in everyday life, such as a thunderstorm or the movement of a skier over a mountain slope, may be the starting point of learning.

Students work on such problems in small tutorial groups of 6 to 10, discussing them initially based only on prior
knowledge. Goal of such initial discussion is to construct a tentative theory explaining the phenomena or events described in the problem-at-hand in terms of its underlying principles or mechanisms. For instance, medical students discussing the fact that Maurice has survived in cold water for so long may come up with the idea that, in cold water, the oxygen needs of the body may be reduced. In addition, they may hypothesize that a child has a smaller body surface and may not decrease in temperature to the same extent as an adult, and so on. The psychology students discussing the car accident may deduce that the mental images produced by the leading question may interfere with the original memory, or that new memories replace old ones the way new files with the same names as old files replace those old files on a hard disk. Alternatively, they may come up with the suggestion that participants in such experiments tend to trust the authority of the experimenter and therefore assume that they have misinterpreted the events in the movie. Hypotheses proposed during this initial analysis of the problem are allowed to be inaccurate, superficial, or outright wrong, as long as they represent the conceptions students hold—or have collaboratively constructed—about the world. It is deemed important that students’ misconceptions are expressed, because this has been demonstrated to facilitate remediation through the confrontation with new, more accurate conceptions (Chinn & Brewer, 1993; Dole & Sinatra, 1998).

Based on this initial discussion, often lasting for more than an hour, learning issues for individual study are formulated. These learning issues usually consist of questions arising from the discussion. For example, the medical students may have formulated issues for self-study such as, What is a mammalian diving reflex? Does the blood circulation to the skin and the extremities shut down? And if so, how? What happens to blood circulation to the vital organs? The psychology students may have come up with issues such as, Why does the specific phrasing of a question influence the response of participants in experiments? Is the original memory changed by the intervening question episode? If yes, how? What are the implications of such experiments for eyewitness testimony?

Students will pursue these learning issues through individual, self-directed learning usually using a variety of resources: books, articles, movies, and Internet sites. These resources may be teacher suggested or student selected, or a combination of both. In most PBL curricula, students are given some responsibility in choosing their own resources. It is suggested that making choices based on one’s own judgment of the importance of a particular source of information supports the experience of being an autonomous learner, a condition conducive to the development of intrinsic interest in the topic at hand. This in turn would foster autonomy (Ryan & Deci, 2000). In particular in the early phase of their studies, however, students tend to have difficulty overseeing their field of study, and therefore, tutor scaffolding takes place here more extensively than in later phases.

After this period of self-directed learning activity, usually lasting for 1 or 2 days, but in some curricula up to a week, students return to their tutorial group, review and share what they have learned, and elaborate upon it. This second meeting is used to explore to what extent the students’ understanding of the problem has developed and whether misconceptions remain that need to be addressed. A further role of continued discussion of the problem is that it enables students to elaborate on the knowledge acquired. It is assumed that elaboration helps students in the long-term retention of the subject matter studied (Pressley et al., 1992; Reder, 1980). Finally, continued discussion of the problem discourages “free riders.” When students know that they are required to share what they have learned, their inclination to let others do the work diminishes. A tutor supervises all these activities. His or her role is not to teach but to guide through conversation and cross-examination (Barrows, 1985).

The process of PBL as described here is embedded in a curriculum. Such curriculum consists of a series of 150 to 350 academic or professional problems that are believed to cover the most important topics of the domain. The curriculum usually is subdivided in modules, each of which is dedicated to a particular theme. In medical education, the themes are often multidisciplinary in nature and related to the major bodily systems such as the cardiovascular system and the gastrointestinal system. In psychology, themes tend to coincide with the various disciplines of the domain. Each of the modules consists of a series of problems, supplemented with a limited number of lectures intended only to clarify difficult issues that remain after students have worked on a particular problem. In addition, curricular time is scheduled for professional and academic skills training. So, in a typical week students would work on two problems, spend 4 to 6 hr in compulsory tutorials, attend 2 hr of lectures, and would be involved in 3 hr of skills training. The remaining time would be free for self-study.

THREE INTERPRETATIONS OF WHAT HAPPENS TO THE LEARNER IN PBL

PBL is not a single educational treatment but a conglomerate of interventions each of which is thought to foster learning. This may be the reason why different authors have tended to define the purpose of PBL differently and stress different aspects of PBL as crucial. We discuss these perspectives here.

Some authors consider PBL to be an example of the information-processing or cognitive constructivist approach to education (Hmelo-Silver, 2004; Norman & Schmidt, 1992; Schmidt, 1993; Schmidt, De Grave, De Volder, Moust, & Patel, 1989; Schwartz & Bransford, 1998). The idea here is that the central goal of PBL is to help students build flexible mental models of the world. The problem represents the part of the world that must be understood, and the small-group discussion and self-study intend to help students
construct a theory explaining the problem in terms of its underlying structure. According to Schmidt (1983), the initial analysis of a problem serves to activate prior knowledge, which is then used to collaboratively construct a tentative model of the situation described. This model is subsequently tested against the available literature and enriched and modified by it. Because the literature is studied with preconceptions activated in mind, discrepancies between faulty prior knowledge and new knowledge can be more easily resolved, and better learning would ensue. In addition, prior knowledge, once activated, would provide better scaffolds for new information. Returning to the problem after individual study serves to further elaborate on what has been learned and to check whether a deeper understanding has evolved. There is some evidence that this is indeed what happens. Discussion of a problem before processing information relevant to that problem facilitates the understanding of that information (De Grave, Schmidt, & Boshuizen, 2001; Schmidt et al., 1989). The effect seems particularly apparent after a delay of several months or even a year (Capon & Kuhn, 2004; Tans, Schmidt, Schade-Hooiveen, & Gijswelaers, 1986). We call curricula that consider PBL as mental model construction Type 1 curricula. Examples of Type 1 curricula are those of Maastricht, The Netherlands (Bouhuijs, Schmidt, Snow, & Wijnen, 1978); Manchester, England (O’Neill, Metcalfe, & David, 1999); Missouri-Columbia, USA (Hoffman, Hosokawa, Blake, Headrick, & Johnson, 2006); and the recent incarnation of the McMaster curriculum (Neville & Norman, 2007).

Other authors, particularly those with ties to medical education, prefer to define PBL as a process of inquiry (Barrows, 1990; Barrows & Tamblyn, 1980; Hmelo, 1998b; Hmelo & Ferrari, 1997). The goal of PBL is here to help students learn the skill of diagnostic reasoning by mimicking the thinking processes of the expert. While working on a problem, the students engage in formulating diagnostic hypotheses, and most of the tutorial time is spent weighing the evidence—signs, symptoms, laboratory data, or physical examination findings—in the light of these diagnostic hypotheses. Problems are often presented to students in a sequential fashion, much in the way they present themselves to experienced doctors: Students are given some facts about a patient (so-called trigger material) and subsequently have to discuss what additional information they would need to solve the problem and why. This information is then provided by the tutor and leads to the formulation of further questions, and so on (Neame, 1989). The role of knowledge acquisition is somewhat blurred in this approach. Although it is acknowledged that students need new information to fix holes in their knowledge base, proponents of this view sometimes seem to suggest that PBL particularly has a role to play in the fostering of “inquiry skills” or “problem-solving skills”—general cognitive skills that can be applied to gather, interpret, and integrate data from any clinical problem. Others have challenged this view, arguing that such skills show little evidence of generalizability or improvement with education (Norman, Patel, & Schmidt, 1990). Some studies seem to support this point of view by demonstrating superiority in diagnostic competence by students from problem-based curricula (Hmelo, 1998a, 1998b; Schuwirth et al., 1999). Curricula that emphasize inquiry or problem solving as a main goal of PBL we call Type 2 curricula. Examples of such curricula are the early McMaster curriculum (Sibley, 1989); the New Mexico curriculum (Kaufman, 1985); and the Newcastle, Australia, curriculum (Neame, 1989).

A third category of authors stresses the fact that students in PBL have (some) autonomy in their choice of resources and are supposed to show agency (Silen & Uhlin, 2008; Toon, 1997). They consider PBL a tool for “learning how to learn.” In this view, knowledge develops so fast as a result of expanding science efforts that, by the time students graduate, much of their knowledge has become obsolete. Therefore, it is more important to acquire skills on how to learn than to learn subject matter. The goal of PBL is to help students acquiring these learning skills; it would prepare them for lifelong learning. There is some evidence that, while in medical school, PBL students make more use of library resources (Blumberg & Michael, 1992; Marshall, Fitzgerald, Busby, & Heaton, 1993; Rankin, 1992). However, there is no evidence to date that any of these learning-related activities can be learned as a skill, nor is there evidence that focusing on learning skills fosters lifelong learning (Juul-Dam, Brunner, Katzenellenbogen, Silverstein, & Christakis, 2001; Ozuah, Curtis, & Stein, 2001; Schmidt, 2000; Shin, Haynes, & Johnston, 1993; Winch, 2008). Curricula that emphasize the acquisition of self-directed learning skills are called Type 3 here. The best-known example is perhaps the Harvard New Pathways curriculum (Tosteson, Adelstein, & Carver, 1994).

**EFFECTS OF PBL: CURRICULUM COMPARISON STUDIES**

The assumptions on which the process of PBL rests, have been studied empirically only to a limited extent (for overviews, see Dolmans & Schmidt, 2006; Hmelo-Silver, 2004; Norman & Schmidt, 1992; Schmidt & Moust, 2000). One reason is that much of the research effort and resources have been focused on curriculum-level outcome studies comparing problem-based with conventional education. Usually, these studies involve populations or large samples of students or graduates from two schools, one of which is labeled “problem-based” and the other “conventional.” Performance on a measure of relevance, predominantly a test of knowledge, is subsequently compared and conclusions are drawn with regard to the effectiveness of PBL as an approach to learning and instruction. Not coincidentally, these full-curriculum comparisons are mainly conducted in the domain of medicine. Medical education is high-stakes education because poor training of medical students can result in
medical errors and, hence, in disabled or even dead patients. In addition, medicine, more than other disciplines, has a tradition of large-scale investigations through which long-term effects of particular medical treatments are studied. From this, it is a small step to apply this approach to educational treatments as well. Finally, PBL requires fairly large initial investments, financial and otherwise, which virtually invites as a first question for research: Is it better than what we did previously?

Comparing different curricula in terms of their outcomes is however somewhat problematic. Most important, unlike what can be done to patients, randomization of students over treatment conditions is almost always impossible. This problem is aggravated by the fact that comparisons are sometimes conducted among schools that are known to admit widely different populations of students, without attempts to equalize students groups post hoc through matching of some sort. Second, treatments cannot be delivered in a blinded fashion as with placebo-drug research. Students and teachers know that they are part of an innovation and may adapt their behavior accordingly (Norman & Schmidt, 2000). Third, neither PBL nor direct-instruction approaches to education are unitary treatments. They consist of many interventions, making it difficult to assign effects to specific parts of the curriculum. Fourth, curricula may call themselves problem based although they fall short in terms of one or more of the six characteristics of PBL mentioned before. Examples are the so-called hybrid problem-based curricula (Bhattacharya, Bhattacharya, & Karmacharya, 2008; Gwee & Tan, 2001; Houlden, Collier, Frid, John, & Pross, 2001). Such curricula employ problems and small-group tutorials combined with extensive direct instruction through lecturing. Fifth, even if curricula pass the six-defining-characteristics-of-PBL test, they may be quite different in their objectives (see the previous discussion about Type 1, 2, and 3 problem-based curricula). And sixth, because medical students are highly selected in terms of the knowledge and skill required to enter medical school, performance on tests are bound to show ceiling effects, leaving little room for improvement. On the other hand, there is no easily available methodological alternative when one is interested in studying the effects of a complex instructional approach to learning. Curriculum comparisons are, therefore, perhaps inevitable means to this end.

META-ANALYSIS OF EFFECTS OF A SINGLE PROBLEM-BASED CURRICULUM

In the next sections, we review the results of a large number of curriculum-comparison studies in which the performance of medical students and graduates of a single, Dutch medical school was compared with the performance of medical students and graduates trained through conventional medical education. There are four reasons why we have chosen to concentrate on effects of a single school rather than to provide a multischool outcome review. The most important reason is that results of multischool outcome studies are inconsistent over PBL curricula. For instance, some U.S. problem-based medical schools are reported to be doing worse on national licensure examinations than conventional comparison schools (e.g., Moore-West & O’Donnell, 1985), others are doing better than their counterparts (Hoffman et al., 2006). Hmelo (1998a, 1998b) was able to demonstrate positive effects of PBL on diagnostic reasoning of medical students in a midwestern and a southern medical school, whereas Patel, Groen, and Norman (1991) found the reverse in two Canadian schools. Other curricular comparisons do not report differences (Albanese & Mitchell, 1993; Berkson, 1993; Collier, 2000; Vernon & Blake, 1993). It is presently unclear why outcomes of apparently similar curricula are not consistently pointing in the same direction. Our hypothesis is that it may have to do with the extent to which, in the various types of schools, knowledge acquisition is valued as a goal in its own right (Type 1, Type 2, Type 3, hybrid). Because of these variations in the way in which PBL is understood and implemented, it is, in our view, for the time being more illuminative to study effects of a single school with a crystallized curriculum of a particular type, invariant over different comparison studies, than to meta-analyze effects of one-shot studies from schools with widely different interpretations of PBL.

A second reason to concentrate on this particular school is that Dutch medical students are admitted to medical schools through a centralized process, in which a weighted lottery procedure is employed based on achievement on a national entrance examination. This procedure (inadvertently) results in groups of students in different schools that are highly similar in terms of past performance, age, gender, socioeconomic status, and motivation to study medicine. For instance, Roelveld (1997) demonstrated that students entering the eight schools between 1989 and 1997 had, every year, almost the same mean high school national examination grade point average. Consequently, comparisons in terms of performance between Dutch medical schools come as close as one can get to controlled field experiments in educational settings. In addition, all schools employ a 6-year curriculum and the subject matter taught is largely overlapping. Again, this facilitates comparisons between different curricula.

Third, the particular school is by far the most extensively studied problem-based curriculum in the world. In this article, that summarizes these studies, we report on 270 comparisons in which students and graduates from this school were involved, more than one third of comparisons published in the literature since the emergence of PBL in higher education (as indicated by a literature search conducted in 2006 available from the first author). A fourth reason to review studies from this particular school was that some of them have been published in outlets not easily accessible, for instance, because they were published in Dutch.

These four reasons led us to focus on effects of this particular curriculum rather than to engage in a meta-analysis...
of results of a wider range of problem-based curricula. To reduce to some extent the generalizability limitations that result from our decision to concentrate on one curriculum, we attempt throughout to put our findings into perspective with the help of results from studies conducted elsewhere.

METHOD

The Medical School Involved

The school that is the focus of our review is the medical school of Maastricht University in The Netherlands. The particular school was founded in 1974, 5 years after the McMaster University medical curriculum was established. All learning starts with problems; tutorial groups form the backbone of the approach, professional skills training is an integral part of the curriculum, and lectures are few. Initially, the curriculum was virtually a copy of the McMaster curriculum, like most other early curricula that pass the six-defining-characteristics-of-PBL test. Like the McMaster curriculum, it emphasized the development of problem-solving skills over subject-matter acquisition and could therefore initially be described as a Type 2 problem-based curriculum. The problem-solving rhetoric, however, led students in the Maastricht curriculum to focus on “playing the doctor;” they would initially tend to study manuals on how to deal with a patient with diabetes (what to ask, how to examine, how to treat) rather than engage themselves in the underlying physiology and biochemistry of the diabetes problem. However, under the influence of studies demonstrating that general clinical reasoning skills cannot really be taught (Barrows, Neufeld, Feighner, & Norman, 1978; Elstein, Shulman, & Sprafka, 1978; Neufeld, Norman, Feighner, & Barrows, 1981), focus in the Maastricht curriculum shifted from an emphasis on problem solving to an emphasis on understanding, from acquisition of clinical-reasoning skills to mental-model construction. Since the late 70s, the curriculum is a robust Type 1 problem-based curriculum with continued emphasis on the acquisition of knowledge for mental-model construction.

There was, initially, another difference between the Maastricht and McMaster curricula. To encourage students to pursue their own learning goals, the McMaster curriculum was largely assessment free. Tutors would provide students with feedback about their performance but achievement tests were absent. In line with its emphasis on knowledge acquisition, the Maastricht curriculum on the other hand introduced a medical knowledge test consisting of 200 to 300 questions covering medicine as a whole. The test is administered to all students, regardless of the year in which these students are studying, four times a year. For each administration of the test, a new version is drawn from a large item bank. Students have to demonstrate progress on each subsequent test, which is why it is referred to as a “progress test” (Van der Vleuten, Verwijnen, & Wijnen, 1996; Wijnen, 1977). More recently, the McMaster curriculum has come around to become a Type 1 curriculum as well (Neville & Norman, 2007), including the introduction of a progress test (Blake et al., 1996). McMaster graduates, initially often scoring below average on National Licensing Examinations in Canada, are now among the better achievers (G. R. Norman, personal communication, December 12, 2008).

The Comparison Schools

The conventional schools involved in almost all of the comparisons to be reported are the other seven Dutch medical schools. In the 70s and 80s, these schools were conventional in the sense that extensive lecturing was the main mode of instruction and learning was largely teacher directed. In the course of the 80s, a number of these schools adapted their curriculum, possibly under the influence of the perceived success of the Maastricht curriculum, cutting down on lectures, introducing small-group work and integrating preclinical with clinical disciplines. Table 1 contains a description of characteristics of these curricula, derived from reports of accreditation committees visiting these schools in the 90s (Visitatiecommissie Geneeskunde en Gezondheidswetenschappen, 1992, 1997).

The table shows that two of the eight medical schools can be characterized as problem-based outright: Maastricht and Groningen (after 1992). Nijmegen (after 1994) comes close, but in this school learning does not start with problems and is more teacher-directed. Comparisons involving the new Groningen and Nijmegen curricula were excluded from the analyses described below.

Selection of Studies Included in the Review

A literature search was initiated using the ISI Web of Science; PsycINFO; Educational Resources Information Center, or ERIC; and PubMed databases. Keywords used were problem-based learning, curriculum, comparison, Maastricht, and all possible combinations and permutations of these terms. In addition, the Google Scholar and Scopus search engines were used to identify additional sources. Next, PBL archives, available on the Internet, were searched. Finally, books, conference proceedings, and conference papers collected at the problem-based school under study were scanned for relevant studies. The initial selection of studies was read and, from these, additional relevant references were identified. Studies were included in which students or graduates from Maastricht medical school were compared with at least one conventional medical school.

Analysis of Findings

For each of the comparisons it was determined which schools were involved. Sometimes, the conventional school was mentioned by name, sometimes it was left anonymous (“School

Downloaded By: [Erasmus University Library / Rotterdamse Leeskabinet / Erasmus MC / Univ Med Centre Rotterdam] At: 11:55 22 December 2009
### TABLE 1
Features of Dutch Medical Curricula Involved in the Meta-Analysis

<table>
<thead>
<tr>
<th>School</th>
<th>Nature of the Curriculum</th>
<th>All Learning Starts With Problems</th>
<th>Students Work in Small Groups</th>
<th>Tutors Do Not Teach Directly</th>
<th>Learning Activities Are Student-initiated</th>
<th>Hours of Lectures per Week</th>
<th>Hours Available for Self-Study per Week</th>
<th>Modular Organization</th>
<th>Integration Between Basic and Clinical Sciences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maastricht</td>
<td>Problem-based</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>3</td>
<td>27</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Leiden</td>
<td>Conventional</td>
<td>No</td>
<td>Partial</td>
<td>No</td>
<td>No</td>
<td>11</td>
<td>20</td>
<td>Partial</td>
<td>Partial</td>
</tr>
<tr>
<td>Amsterdam UvA</td>
<td>Conventional</td>
<td>No</td>
<td>Partial</td>
<td>No</td>
<td>No</td>
<td>8</td>
<td>21</td>
<td>Partial</td>
<td>Partial</td>
</tr>
<tr>
<td>Amsterdam VU</td>
<td>Conventional</td>
<td>No</td>
<td>Partial</td>
<td>No</td>
<td>No</td>
<td>10</td>
<td>22</td>
<td>Partial</td>
<td>Partial</td>
</tr>
<tr>
<td>Rotterdam</td>
<td>Conventional</td>
<td>No</td>
<td>Partial</td>
<td>No</td>
<td>No</td>
<td>12</td>
<td>20</td>
<td>Partial</td>
<td>Partial</td>
</tr>
<tr>
<td>Utrecht</td>
<td>Conventional</td>
<td>No</td>
<td>Partial</td>
<td>No</td>
<td>No</td>
<td>11</td>
<td>16</td>
<td>No</td>
<td>Partial</td>
</tr>
<tr>
<td>Groningen until 1992</td>
<td>Conventional</td>
<td>No</td>
<td>No</td>
<td>—</td>
<td>No</td>
<td>12</td>
<td>18</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Nijmegen until 1994</td>
<td>Conventional</td>
<td>No</td>
<td>No</td>
<td>—</td>
<td>No</td>
<td>14</td>
<td>20</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Groningen after 1992</td>
<td>Problem-based</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>5</td>
<td>27</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Nijmegen after 1994</td>
<td>Active learning, non-problem-based</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>2</td>
<td>28</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

**Note.** Adapted from Schmidt, Cohen-Schotanus, and Arends (2009). Numbers of hours per week are averaged over the first 4 preclinical years. Modular organization of the curriculum implies that subject matter is offered in a sequential rather than in a parallel fashion. Integration between basic and clinical sciences refers to the fact that these curricula present the disciplines of medicine in an integrated fashion.
Medical Knowledge

As has been described previously, the problem-based school under study has developed a medical knowledge test consisting of 200 to 300 questions covering medicine as a whole, which all students routinely take four times a year. The test is called “progress test” because it helps monitoring the students’ advancement through the program. It has high reliability and has been demonstrated to have high validity as well (Van der Vleuten et al., 1996).

The first comparisons using the progress test were conducted in the late 70s and early 80s but published later (Verwijnen et al., 1990). These comparisons concerned studies involving all students from the problem-based school, and large groups of volunteers from three conventional medical schools, indicated as A, B, and C. The tests also included questions that had been submitted by the other medical schools involved in the comparison study. As the Appendix demonstrates, many of these early comparisons slightly favor the conventional schools involved, resulting in a negative overall effect of \( d = -0.16 \). Using the progress test, Verhoeven et al. (1998) conducted comparison studies between all the students of the problem-based school and the majority of the students from the medical school of Nijmegen University. An overall small \( d \) of \(-0.13\) emerged from these comparisons. A series of studies by Albano and colleagues (1996) compared level of medical knowledge of the Dutch problem-based school with four conventional medical schools from Germany and Italy. Overall medium effect size was \( 0.46 \). Presently, the progress test is administered routinely four times a year to all students of four out of eight medical schools in The Netherlands. All schools contribute items to the test (Van der Vleuten et al., 2004). One of these schools is a conventional school. In the Appendix, comparisons carried out in the course of 2006 are displayed. The overall effect size for the comparisons during that particular year was \( 0.05 \).

Using progress test data, a number of comparative studies were conducted on knowledge of anatomy, one of the subdisciplines of medicine (Imbos, Drukker, Van Mameren, & Verwijnen, 1984; Prince et al., 2003). The assumption underlying these studies was that students studying in a problem-based medical curriculum would perhaps know less about basic-science subjects (such as anatomy or biochemistry) than their colleagues trained in more conventional centers, because they were expected to concentrate on applied, clinical problems earlier in their study. In the Imbos and colleagues (1984) study, students from the problem-based school started at lower levels of anatomy knowledge than randomly selected students from the conventional comparison school, but unlike the latter, the PBL students’ knowledge continued to grow while in the clerkships. The study does not report means and standard deviations. However, Figures 5 and 6 in their contribution suggest no differences between groups. The Prince et al. study employed 4th-year students sampled from six conventional medical schools. Mean overall effect size was \( d = \).
.11. Of interest, a medium positive effect of PBL was found on items presented in the context of clinical cases \( (d = .29) \), whereas no overall effect was found on items presented in isolation \( (d = .02) \). Finally, Schmidt, Vermeulen, and Van der Molen (2006) asked 1,400 graduates from two schools to rate themselves with regard to their level of medical knowledge. The effect size \( d \) was equal to \(-.18\).

Excluding these self-ratings and the international comparisons (Albano et al., 1996), the overall weighted effect size averaged over the 90 comparisons involving the problem-based school and various Dutch conventional schools was equal to \( d = .07 \), signifying a small positive effect.

### Diagnostic Reasoning

A primary task of the competent physician is not so much to recall medical knowledge but to use that knowledge in the diagnosis of medical problems. Typically, research on medical diagnosis presents a number of cases to participants of different levels of expertise and requires them to produce a (differential) diagnosis. Relatively little research using this methodology has been conducted to date on whether students who attended a problem-based medical school are better at solving diagnostic problems than students attending conventional training. Worldwide, only studies by Hmelo (1998a, 1998b) and Patel and colleagues (1991) are available. These studies show mixed results.

Several studies were conducted in The Netherlands. In one of these, volunteers from the problem-based school, and volunteers from the conventional schools in Groningen, and Amsterdam were tested using an assessment instrument intended to evaluate diagnostic reasoning. The instrument consisted of 30 case descriptions covering all body systems and being epidemiologically representative (Schmidt et al., 1996). Participants wrote up a differential diagnosis for each of the 30 cases, and the number of accurate diagnoses was counted. Mean effect size was equal to \(-.08\). In a similar second study, samples of students from the problem-based school and Groningen were presented 60 short case descriptions (Schuwirth et al., 1999). In this study, the students from the problem-based school achieved significantly higher scores \( (d = .38) \).

Earlier, Boshuizen and her collaborators carried out two comparisons focusing on the relation between knowledge and reasoning. Boshuizen, Schmidt, and Wassmer (1994) compared the performances of small samples of students from Maastricht and Amsterdam. These (preclinical) students were asked to explain how a specific metabolic deficiency and a specific disease could be related, for example, “How does a genetic deficiency of pyruvate kinase lead to hemolytic anemia?” In answering such question, knowledge about biochemistry and about internal medicine must be applied and integrated. Students from the problem-based curriculum appeared to take an analytical approach to the problem by first exploring the biochemical aspects of the problem, later linking them to clinical aspects. Students in the conventional curriculum tended toward a more memory-based approach. This latter strategy, however, resulted in significantly fewer accurate answers and more failures by the students from the conventional program. These results suggested that students in a problem-based curriculum better integrate their knowledge than students in a traditional curriculum, resulting in more accurate reasoning. This finding concurs with findings by Hmelo (1998a, 1998b). Another study using a clinical case recall paradigm (recall is often seen as a measure of expertise since De Groot’s seminal studies in chess) and measuring processing speed (shorter processing is considered a sign of better understanding) demonstrated the Maastricht students to be better (Boshuizen & Claessens, 1982).

In summary, PBL seemed to have a small positive impact on the diagnostic competencies of students \( (overall \ d = .11) \). As suggested by the Boshuizen studies, this effect is possibly caused by a somewhat better integration of biomedical knowledge into clinical knowledge.

### Communication Skills and Other Work-Supporting Competencies

Communication skills and other work-supporting competencies such as the ability to work efficiently and to work in teams are considered to be important for every professional. Van Dalen et al. (2002) presented students from Maastricht and Leiden medical schools with four cases. Blinded observers judged their level of communication skills. The participants of the problem-based school displayed a higher level of skill in this respect. A large effect size of 1.46 was reported.

In a study among a large group of physicians who graduated from the medical schools in Maastricht and Rotterdam, participants were asked to assess their own skill levels with regard to eighteen professional competencies (Schmidt et al., 2006). The oldest participants had graduated 19 years previously; the mean time passed since graduation was slightly more than 9 years. The competencies concerned interpersonal skills such as the ability to collaborate with others, cognitive skills such as the ability to find relevant information quickly, general academic skills such as the ability to conduct scientific research or write a report, and efficiency skills such as the ability to comply with deadlines or to work under pressure. Effects showed the problem-based school to be superior. The overall weighted effect size was equal to \(.69\). The effect of PBL was largest in the area of interpersonal skills \( (Md = 1.39) \). It was attributed in part to the fact that students in a problem-based curriculum have to work extensively with others in small groups, so they have more opportunities to practice those interpersonal skills.

### Medical Skills

In addition to the general professional skills important to any type of professional working with people, medical
students also have to acquire the more domain-specific practical medical skills, such as blood pressure measurement, abdominal examination, or resuscitation. In the Schmidt et al. (2006) study, graduates from the PBL curriculum indicated that they had mastered such medical skills at a higher level than colleagues who had graduated from the other university. This finding concurs with findings of several other studies (Remmen et al., 1999; Remmen et al., 2001; Scherpnier, 1997). One such study compared two groups of PBL students actually performing a number of medical skills in comparison with Groningen students (Scherpbier, 1997). The PBL students showed higher levels of achievement. Five studies tested knowledge about skills (a variable closely related to actual skill level: Van der Vleuten, Van Luyk, & Beckers, 1989) among volunteers from the medical schools of Maastricht, Groningen, Antwerp, and Ghent—the latter two being Belgian schools (Remmen et al., 1999; Remmen et al., 2001; Scherpnier, 1997). On this test, the Dutch students achieved significantly higher marks than the Belgian students. Moreover, the PBL students outperformed Groningen students, particularly in the last 2 years of medical school. The overall weighted effect size $d$ for the level of mastery of these skills, excluding the comparisons with the Belgian schools, was equal to .83.

Student and Expert Judgments of the Quality of Problem-Based Instruction

Each year, 20,000 Dutch students are polled about the quality of their education for a Consumer Reports–type guide to higher education, called the “Keuzegids hoger onderwijs” (Steenkamp et al., 2006; Steenkamp et al., 2004; Steenkamp et al., 2008). In this guide, programs within a particular domain are compared nationwide to help prospective students make more informed career and education decisions. For each program, randomly selected students receive a questionnaire asking them to rate their curriculum on 10 dimensions: (a) the quality of the content of the curriculum, (b) electives, (c) coherence of the curriculum, (d) the instructional approach, (e) preparation for professional practice, (f) the quality of the teachers, (g) communication with the faculty, (h) the extent to which the program can be completed within the timeframe available, (i) the quality of the classrooms, and (j) the student facilities. The weighted effect size of the mean was equal to .66, signifying a medium effect. In each edition of the student guide since its inception in the early 90s, the problem-based school came out best. Students particularly like the instructional approach emphasizing independent study and critical thinking; here the difference with the other schools is largest. In addition, the problem-based curriculum rates high with regard to its contents, its coherence, and its preparation for the profession.

In 5-year cycles, an accreditation committee of national experts surveys all medical schools. The resulting reports are used for decision making at the national level. The findings of these experts, published over the last 20 years, generally concur with the results of the student surveys (e.g., Looijenga, 2004).

Graduation Rates and Study Duration

Graduation rates (i.e., the percentage of students succeeding of those initially entering a program) and study duration data are collected systematically in The Netherlands since 1989, but some national data are available from medical schools for previous years. Remember that the first cohorts of Maastricht medical students entered medical school in 1974 and began to graduate in 1980. A study by Post, De Graaff, and Drop (1986) provides data relevant for this period. They gathered graduation rates for all medical students entering the seven medical schools in The Netherlands in 1970 and compared their performance with the four cohorts entering Maastricht medical school between 1974 and 1977. The authors showed that, after 6 years, 64% of the first cohort from Maastricht medical school had acquired its medical degree, whereas 0% of those of the other medical schools had received a degree by that time. Nine years after entering medical school, 96% of the problem-based cohort had acquired its degree, whereas the graduation percentage for the other schools was 64%.

More recently, a study was conducted on data collected from all students entering Dutch medical education between 1989 and 1998 (Schmidt, Cohen-Schotanus, & Arends, 2009). This study demonstrated that the problem-based curriculum had a medium effect on both graduation rate (overall $d = .33$) and study duration ($d = -.68$). The mean difference in graduation rate 9 years after admission was 12%; the mean difference in study duration was .54 year. Not only does the problem-based school have a considerable lower dropout than the other schools, students receive their degree faster as well.

**DISCUSSION**

The review presented here was conducted to provide a comprehensive overview of the effects of one problem-based curriculum on its students and graduates. The assumption underlying this research was that it might be more illuminating, at this stage of development of PBL, to summarize effects of one well-established school with a crystallized curriculum than to add up sometimes conflicting findings from schools that stress widely different objectives. In addition, the particular school under study is the most researched curriculum worldwide, and the students and graduates from the different schools involved in the comparisons were known to be highly similar in terms of prior performance and homogeneous in term of previous learning experiences when they enter medical school (Roelleveld, 1997). To that end, 270 comparisons were assessed. For most of these comparisons we were able to compute effect sizes. In addition, within the different...
dimensions of performance in which we were interested, overall effect sizes were computed weighted according to the size of the populations or samples involved. The results support the following conclusions.

First, effects of PBL are strongest in the professional skills domain. The effect was particularly strong in the area of interpersonal skills. In this domain, the average student from the PBL curriculum leaves behind around 92% of the students from the conventional schools (effect sizes can be interpreted as z scores; an effect size of 1.39 has an associated p value of .9177). This finding is in agreement with findings from other problem-based schools (Mennin, Kalishman, Friedman, Pathak, & Snyder, 1996; Santos Gomez, Kalishman, Rezler, Skipper, & Mennin, 1990; Woodward & McAuley, 1983). It suggests that collaborating in small tutorial groups indeed facilitates the acquisition of such skills. In addition, students in the PBL curriculum perform consistently better with regard to practical medical skills. Here, the average PBL student surpasses 79% of his or her colleagues in conventional schools. The skills training program at the particular school is tightly integrated with the theoretical parts of the curriculum (i.e., students are trained in abdominal examination skills in the context of the gastrointestinal system module). In addition, these skills are practiced throughout the curriculum, rather than only prior to, or during, clinical rotations. Some suggest that these are reasons why students do so much better in this domain (Scherpbier, 1997).

Second, the educational approach represented by PBL seems to be highly appreciated. The problem-based curriculum under study came consistently out first in national comparisons in which students rated the program they were in on 10 dimensions of instructional quality. In this domain, the average PBL student leaves behind 75% of his colleagues. Tellingly, Groningen medical school, a relatively new problem-based curriculum in The Netherlands, often comes consistently out first in national comparisons. In this domain, the average student in a conventional curriculum were 3% and 5%, respectively. So, students studying the problem-based curriculum did not know much more about medicine than those studying in conventional schools and their diagnostic competencies did not differ much. Because diagnostic reasoning is a knowledge-based activity, the latter result may not be surprising. However, our findings and those of others (Albanese & Mitchell, 1993; Colliver, 2000; Dochy et al., 2003; Vernon & Blake, 1993) do not seem to fit results of classroom-level experiments studying the effect of PBL on knowledge acquisition (Capon & Kuhn, 2004; De Grave et al., 2001; Schmidt et al., 1989). These studies demonstrate positive and sizable effects of problem discussion on the comprehension of subsequently presented information. In an attempt to resolve this apparent paradox, Schmidt, Cohen-Schotanus, and Arends (2009) called attention to the fact that the differences in graduation rate demonstrate that groups of students involved in curriculum comparisons studies may only partially be similar. They argued that these studies may be biased to some extent against the problem-based curriculum because the latter curriculum retain students that would have been dropped out of any conventional curriculum. This selection bias would actually mask effects of PBL on knowledge acquisition studies in curriculum comparison studies.

The studies just summarized may also have suffered to some extent from the fact that in many of the comparisons
carried out, populations of medical students from Maastricht were compared to samples of volunteers from other universities. Those who conducted the comparisons did not perceive this as a serious problem and, therefore, did generally not control for it. However, volunteers tend to be the better students (Rosenthal & Rosnow, 1975). Where it was possible to check this, the volunteers indeed proved to be the slightly better students (Prince et al., 2003; Verhoeven et al., 1998). Consequently, the results of a number of the studies may be somewhat biased against the effects of PBL. This may in particular be the case for the studies that tested knowledge. On the other hand, in the studies involving entire student populations on both sides, students from the problem-based program did do slightly better, but the effect was small (Van der Vleuten et al., 2004).

The studies reviewed in this article were conducted involving only one problem-based medical school, the medical school of Maastricht University in The Netherlands. Therefore, generalizability of the findings presented here can only be claimed to the extent that these findings have been corroborated in studies from other schools, which has happened to some extent. We have already noted that Groningen medical school, the second problem-based school in The Netherlands, has demonstrated similar effects on graduation rates and on student appreciation. Two problem-based psychology curricula in the same country also demonstrated higher graduation rates and higher student preference rates than their counterparts (Severiens & Schmidt, 2009; Steenkamp et al., 2006).

POSTSCRIPT: CONDITIONS UNDER WHICH CONSTRUCTIVIST CURRICULA MAY SUCCEED

Recently, PBL has come under attack for various reasons (Colliver, 2000; Kirschner et al., 2006; Shanley, 2007). The most principled of these is the critique by Kirschner et al. (2006). Kirschner and colleagues consider PBL, like other constructivist approaches, a form of minimally guided instruction. Because, according to these authors, minimally guided instruction is incompatible with the cognitive architecture of the human mind, the outcomes of such education are bound to be suboptimal compared with approaches in which more direct instruction is provided. They hypothesize, therefore, that learning of students in curricula based on constructivist ideas must be poorer than learning in curricula using direct instruction.

The findings reported in this article do not support their hypothesis. There is no indication in our findings that students in a problem-based curriculum learn less content knowledge. In addition, PBL seems to promote, more than conventional education does, the development of professional skills. It does this in ways preferred by students. In addition, it does this in a more effective and efficient way, as witnessed by the higher graduation rates and lesser delays. The question is then, Why is this so? As discussed in the Introduction section, PBL curricula tend to do with much less direct instruction than conventional programs. Does this imply that direct instruction is perhaps less necessary for learning to take place than Kirschner and colleagues believe?

At his point, it may be appropriate to briefly ponder about what teachers generally do to help students learn. The teacher’s repertoire of activities to support student learning may be considered to consist of at least six categories: (a) transmitting subject matter, (b) providing explanations, (c) directing students toward what is important to learn, (d) engaging in pedagogically purposeful conversations with students, (e) encouraging students and rewarding them, and (f) providing feedback upon learning. The first three activities usually take place through some form of lecturing and can be considered direct instruction. The last three are largely responses to, and dependent on, the students’ actual learning activities in- and outside the classroom. They can therefore be considered instructional scaffolding. Although PBL is relatively poor in direct instruction, it is rich in instructional scaffolding (Schmidt, Loyens, Van Gog, & Paas, 2007). This relative scarcity of direct instruction, however, does not imply that providing explanations and giving directions are absent in PBL (transmission is absent), but they occur only in response to needs expressed by students and are by no means the exclusive domain of the teacher in front of the class. In fact, in PBL these functions are taken over by the other elements of the instructional arrangement: the problems, the peers, and the tutor. Self-directed learning is an additional important component. However, these pedagogical tools of constructivist education only contribute to learning if certain conditions are satisfied. These conditions are briefly discussed in the remainder of this section.

Roles of Problems in PBL

From a cognitive perspective, problems are puzzles that, by their enigmatic nature, evoke epistemic curiosity in students (Inagaki & Hatano, 1977). Curiosity can be interpreted as a form of cognitively induced deprivation that arises from the perception of a gap in knowledge or understanding; learners therefore engage in activities closing that gap (Loewenstein, 1994). Several studies have demonstrated increased levels of curiosity in students after being confronted with a problem in their domain of study. These feelings of curiosity were largely maintained throughout the PBL cycle and were related to subsequent achievement (De Volder, Schmidt, Moust, & De Grave, 1986; Rotgans, 2009). Second, to the extent that problems can be perceived as pertinent to one’s future profession, they contribute to the experience of relevance of the curriculum. It may, therefore, come as no surprise that PBL students report higher levels of perceived relevance of their education in the quality-of-instruction surveys (Steenkamp et al., 2006; Steenkamp et al., 2004; Steenkamp et al., 2008). Third, problems constitute important coordinating elements
in the curriculum. Small-group discussions, subject matter to be studied, lectures, and skills training are all centered on the problem at hand. This may support organizing knowledge in memory such that it can be more easily retrieved and used.

Effects of Small-Group Learning

The cognitive benefits of small-group learning have been extensively documented in the literature (e.g., Slavin, 1996; Springer, Stanne, & Donovan, 1999; Webb, Troper, & Fall, 1995), and need no further discussion here. For cognitive effects of small groups in PBL and their interpretation, the reader is referred to Capon and Kuhn (2004), De Grave et al. (2001), and Schmidt et al. (2007). Here, we would like to draw attention to two other functions of small-group learning. First, the tutorial group is a source of friendships. In addition, the group enables students to develop more personal relationships with teachers than is possible in the larger classroom. Both factors are considered to be protective against premature dropout (Tinto, 1997). Second, regular small-group tutorials in problem-based schools provide peer pressure and natural deadlines for work to be completed and, therefore, encourage students not to postpone self-study. In our view, this is a reason why students in the PBL curriculum graduate earlier than students from conventional schools, as documented in this review (Schmidt et al., 2009).

Role of the Tutor

The tutor’s contribution to learning in PBL curricula has been subject of considerable debate. This debate centered on the question of whether a tutor needs subject-matter knowledge to guide students or whether the possession of appropriate facilitation skills is sufficient. A number of studies have demonstrated effects of tutor expertise on their students’ achievement and effort. Davis, Nairn, Paine, Anderson, and Oh (1992), for instance, showed that student performance on a test measuring knowledge of influenza was enhanced when their tutors entertained an active research interest in that field. Eagle, Harasym, and Mandin (1992) demonstrated that students guided by content-expert tutors produced more than twice as many learning issues for self-directed learning and spent almost twice the amount of time on self-study. Schmidt, Vanderarend, Moust, Kokx, and Boon (1993) found similar effects of subject-matter expertise on achievement. Other studies, however, failed to demonstrate noticeable effects (DesMarchais & Black, 1991; Swanson, Stalenhoef-Halling, & Van der Vleuten, 1990). One hypothesis explaining this discrepancy is that subject-matter expertise of the tutor seems to play a role predominantly when the scaffolds provided by the learning environment itself: the problems, the resources, do not contain sufficient cues as to what is important to study. Under such circumstances, students seem to rely on their tutor for guidance and might profit if their tutor happens to be someone knowledgeable regarding the subject under study (Schmidt, 1994).

Effects of Self-Directed Learning

Self-directed learning is seen as a core element of PBL, and much emphasis is put on the developing ability of students to regulate their own learning. It should be noted here that the various instructional activities conducted both in the problem-based and in the conventional context are intended to support such self-study of students. However, emphasizing instructional scaffolding in the curriculum over direct instruction causes students to spend their time differently. As Table 1 illustrates, students in the problem-based curricula spend less time in lectures and have more time for self-study. In the problem-based school under study, students received three hours of lectures each week and had 27 hr available for self-study, whereas teachers in the conventional schools lectured on average 14 hr, leaving room for 19 hr of self-study. As a result of the emphasis on self-directed learning, PBL students were shown to borrow more books and seek for more learning resources in the library than their counterparts in conventional schools (Blumberg & Michael, 1992; Marshall et al., 1993; Rankin, 1992), suggesting that these students are more self-reliant and feel more responsible for their own learning, a hypothesis corroborated by the quality-of-instruction surveys (Steenkamp et al., 2006; Steenkamp et al., 2004; Steenkamp et al., 2008). A recent study involving graduation-rate and study-duration data from almost 14,000 graduates of the eight medical schools in The Netherlands found that time available for self-study in the various curricula was correlated .44 with graduation rate and -.48 with study duration; that is, the more time for self-study was available in these curricula, the more students graduated and the faster they were able to do this (Schmidt, Cohen-Schotanus, Van der Molen, et al., 2009). These findings suggest that time available for self-study, as afforded in PBL, leads to better achievement of more students, hence to fewer delays and higher graduation rates. The study also demonstrated that the more lectures these students received, the fewer students graduated and the longer it took them to graduate. This finding is ironic in the light of the recent emphasis in the literature on the importance of direct instruction, because it demonstrates that this kind of support is not necessarily good for students.

So, in contrast to what is suggested in recent literature, constructivist learning, as epitomized by the problem-based curriculum studied in this meta-analysis, does work. It does work even while teachers concentrate on providing instructional scaffolding rather than on direct instruction. We suggest, however, that constructivist curricula may succeed only if four conditions are met. First, problems or assignments used as the starting point of small-group discussion and self-directed learning should be promoting epistemic curiosity and should be perceived by students as relevant to their
personal strivings. Second, small-group work should enable the activation of prior knowledge and elaboration on what is learned. Third, tutors should engage themselves actively in didactic conversations with the learners and provide appropriate scaffolds. Fourth, students need ample time for self-directed learning using resources that (to some extent) represent their own interests and preferences. However, this self-directed learning should not be unrestricted; knowledge acquired as a result of this learning should be checked. The peer pressure provided by frequent tutorials and the natural deadlines for work to be completed that these meetings afford seem to encourage students not to postpone self-study and, hence, are helpful in this respect.

For the time being, these conclusions apply only to curricula that help students acquire knowledge for mental model construction (Type 1 curricula), such as the problem-based curriculum that was the subject of the meta-analysis and the other (Dutch) curricula for which similar findings were reported. Whether they also apply to curricula that consider inquiry or problem solving as their focus of attention remains to be demonstrated.

REFERENCES


### APPENDIX

Summary of Studies Comparing, on Various Outcome Measures, the Performance of Students and Graduates of the Problem-Based Medical Curriculum Under Study and Conventional Medical Schools

<table>
<thead>
<tr>
<th>Author</th>
<th>Conventional Comparison School</th>
<th>Sample Information (Population/Volunteers/Random Selection, N)</th>
<th>No. of Classes Involved and Level of Training of Participants</th>
<th>Measuring Instrument Information</th>
<th>Effect Sizea</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verwijnen et al., 1990</td>
<td>School A</td>
<td>PBL-school: Population (N = 266)</td>
<td>6 classes: 1st–6th year</td>
<td>64 MC-items randomly stratified selected from Progress Testb</td>
<td>.44</td>
</tr>
<tr>
<td></td>
<td>School B</td>
<td>Conventional school: Volunteers (N = 550)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verwijnen et al., 1990</td>
<td>School A</td>
<td>PBL-school: Population (N = 266)</td>
<td>6 classes: 1st–6th year</td>
<td>64 MC-items randomly stratified selected from Progress Testb</td>
<td>−.10</td>
</tr>
<tr>
<td></td>
<td>School B</td>
<td>Conventional school: Volunteers (N = 703)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verwijnen et al., 1990</td>
<td>School A</td>
<td>PBL-school: Population (N = 174)</td>
<td>3 classes: 4th–6th year</td>
<td>70 MC-items randomly stratified selected from Progress Test</td>
<td>−.01</td>
</tr>
<tr>
<td></td>
<td>School B</td>
<td>Conventional school: Volunteers (N = 196)</td>
<td></td>
<td></td>
<td>−1.02</td>
</tr>
<tr>
<td>Verwijnen et al., 1990</td>
<td>School A</td>
<td>PBL-school: Population (N = 189)</td>
<td>3 classes: 4th–6th year</td>
<td>64 MC-items randomly stratified selected from Progress Test</td>
<td>.16</td>
</tr>
<tr>
<td></td>
<td>School B</td>
<td>Conventional school: Volunteers (N = 390)</td>
<td></td>
<td></td>
<td>−.06</td>
</tr>
<tr>
<td>Verwijnen et al., 1990</td>
<td>School C</td>
<td>PBL-school: Population (N = 565)</td>
<td>6 classes: 1st–6th year</td>
<td>64 MC-items randomly stratified selected from Progress Test</td>
<td>−.14</td>
</tr>
<tr>
<td></td>
<td>School C</td>
<td>Conventional school: Volunteers (N = 677)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verwijnen et al., 1990</td>
<td>School C</td>
<td>PBL-school: Population (N = 565)</td>
<td>6 classes: 1st–6th year</td>
<td>64 MC-items randomly stratified selected from Progress Test</td>
<td>−.15</td>
</tr>
<tr>
<td>Verhoeven et al., 1998</td>
<td>Non-PBL School Q</td>
<td>PBL-school: Population (N = 181)</td>
<td>3 classes: 1st, 4th, and 6th year</td>
<td>Progress Test: 198 items</td>
<td>.46</td>
</tr>
<tr>
<td>Verhoeven et al., 1998</td>
<td>Non-PBL School R</td>
<td>PBL-school: Population (N = 811)</td>
<td>6 classes: 1st–6th year</td>
<td>Progress Test: 198 items</td>
<td>.61</td>
</tr>
<tr>
<td>School</td>
<td>Type</td>
<td>Sample Size</td>
<td>Class Years</td>
<td>Progress Test</td>
<td>Score</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------</td>
<td>------------------------------</td>
<td>-----------------</td>
<td>---------------</td>
<td>--------</td>
</tr>
<tr>
<td>S</td>
<td>Conventional school: Random sample ($N = 339$)</td>
<td>6 classes: 1st–6th year</td>
<td>Progress Test: 198 items</td>
<td>.55</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>PBL-school: Population ($N = 811$)</td>
<td>5 classes: 1st–3rd, 5th and 6th year</td>
<td>Progress Test: 198 items</td>
<td>.78</td>
<td></td>
</tr>
<tr>
<td>Van der Vleuten et al., 2004</td>
<td>Nondisclosed medical school</td>
<td>6 classes: 1st–6th year</td>
<td>Progress Test: 187 items</td>
<td>.06</td>
<td></td>
</tr>
<tr>
<td>Test October 2006</td>
<td>Conventional school: Population ($N = 1,420$)</td>
<td>6 classes: 1st–6th year</td>
<td>Progress Test: 189 items</td>
<td>-.08</td>
<td></td>
</tr>
<tr>
<td>Test December 2006</td>
<td>Conventional school: Population ($N = 1,350$)</td>
<td>6 classes: 1st–6th year</td>
<td>Progress Test: 191 items</td>
<td>.09</td>
<td></td>
</tr>
<tr>
<td>Test February 2007</td>
<td>Conventional school: Population ($N = 911$)</td>
<td>6 classes: 1st–6th year</td>
<td>Progress Test: 194 items</td>
<td>.18</td>
<td></td>
</tr>
<tr>
<td>Test May 2007</td>
<td>Conventional school: Population ($N = 952$)</td>
<td>6 classes: 1st–6th year</td>
<td>Anatomy items from several Progress Tests</td>
<td>See text</td>
<td></td>
</tr>
<tr>
<td>Imbos et al., 1984</td>
<td>Nondisclosed medical school</td>
<td>6 classes: 1st–6th year</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prince et al., 2003</td>
<td>School C</td>
<td>PBL-school: Random sample of volunteers' ($N = 75$)</td>
<td>4th year</td>
<td>Anatomy test of 138 items presented in the clinical context of 16 cases</td>
<td>.65</td>
</tr>
<tr>
<td>Prince et al., 2003</td>
<td>School E</td>
<td>PBL-school: Random sample of volunteers ($N = 75$)</td>
<td>4th year</td>
<td>Anatomy test consisting of 48 items without clinical context</td>
<td>-.16</td>
</tr>
<tr>
<td>Prince et al., 2003</td>
<td>School F</td>
<td>PBL-school: Random sample of volunteers ($N = 75$)</td>
<td>4th year</td>
<td>Anatomy test consisting of 48 items without clinical context</td>
<td>.05</td>
</tr>
</tbody>
</table>

(Continued on next page)
<table>
<thead>
<tr>
<th>Author</th>
<th>Conventional School</th>
<th>Sample Information</th>
<th>No. of Classes Involved and Level of Training of Participants</th>
<th>Measuring Instrument Information</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prince et al., 2003</td>
<td>School G</td>
<td>PBL-school: Random sample of volunteers ($N = 75$)</td>
<td>4th year</td>
<td>Anatomy test consisting of 48 items without clinical context</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Conventional school: Random sample of volunteers ($N = 33$)</td>
<td></td>
<td>Anatomy test of 138 items presented in the clinical context of 16 cases</td>
<td>$-.55$</td>
</tr>
<tr>
<td>Prince et al., 2003</td>
<td>School H</td>
<td>PBL-school: Random sample of volunteers ($N = 75$)</td>
<td>4th year</td>
<td>Anatomy test consisting of 48 items without clinical context</td>
<td>$-.03$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Conventional school: Random sample of volunteers ($N = 38$)</td>
<td></td>
<td>Anatomy test of 138 items presented in the clinical context of 16 cases</td>
<td>$ .51$</td>
</tr>
<tr>
<td>Schmidt et al., 2006</td>
<td>Conventional School</td>
<td>PBL-school: Volunteers ($N = 820$)</td>
<td>graduates</td>
<td>Anatomy test consisting of 48 items without clinical context</td>
<td>$ .21$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Conventional school: Volunteers ($N = 621$)</td>
<td></td>
<td>Self-report measure of level of medical knowledge</td>
<td>$-.18$</td>
</tr>
<tr>
<td>Diagnostic reasoning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boshuizen &amp; Claessen, 1982</td>
<td>Utrecht</td>
<td>PBL-school: Volunteers ($N = 21$)</td>
<td>3 classes: 3rd–5th year</td>
<td>Propositions recalled from 2 clinical cases</td>
<td>$ .50$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Conventional school: Volunteers ($N = 26$)</td>
<td></td>
<td>Time needed to process 2 clinical cases</td>
<td>$-.56^d$</td>
</tr>
<tr>
<td>Boshuizen et al., 1994</td>
<td>Conventional School</td>
<td>PBL-school: Volunteers ($N = 21$)</td>
<td>3 classes: 3rd–5th year</td>
<td>One clinical problem: thinking aloud. Verbatim transcript analyzed in terms of quality and accuracy of answers</td>
<td>See text</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Conventional school: Volunteers ($N = 26$)</td>
<td></td>
<td>Providing the right diagnosis for 30 clinical cases</td>
<td>$-.36$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Conventional school: Volunteers ($N = 4$)</td>
<td>2 classes: 3rd–4th year</td>
<td>Providing the right diagnosis for 30 clinical cases</td>
<td>$ .20$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Conventional school: Volunteers ($N = 200$)</td>
<td>5 classes: 2nd–6th year</td>
<td>Case-based computerized test: 60 cases</td>
<td>$ .38$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Groningen PBL-school: Volunteers ($N = 200$)</td>
<td></td>
<td>Providing the right diagnosis for 30 clinical cases</td>
<td>See text</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Groningen PBL-school: Volunteers ($N = 78$)</td>
<td></td>
<td>Case-based computerized test: 60 cases</td>
<td>See text</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Groningen PBL-school: Volunteers ($N = 177$)</td>
<td></td>
<td>Case-based computerized test: 60 cases</td>
<td>See text</td>
</tr>
<tr>
<td>Communication skills and other work-supporting competencies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Van Dalen et al., 2002</td>
<td>Leiden</td>
<td>PBL-school: Random sample ($N = 76$)</td>
<td>2 classes: 4th and 6th year</td>
<td>Multiple station examination: 4 cases; ratings of communication skills by blinded observers</td>
<td>$1.46$</td>
</tr>
<tr>
<td>Schmidt et al., 2006</td>
<td>Conventional School</td>
<td>PBL-school: Random sample ($N = 57$)</td>
<td>Graduates</td>
<td>Self-report measure of level of communication skills</td>
<td>$1.30$</td>
</tr>
<tr>
<td>Study</td>
<td>School Type</td>
<td>Participants</td>
<td>Method</td>
<td>Results</td>
<td></td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>------------------------------</td>
<td>--------------</td>
<td>----------------------------------------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>Schmidt et al., 2006</td>
<td>Conventional School</td>
<td>Volunteers (N = 621)</td>
<td>Self-report questionnaire of level of cognitive strategies</td>
<td>.78</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PBL-school: Volunteers (N = 820)</td>
<td>Volunteers (N = 820)</td>
<td>Self-report of level of academic skills</td>
<td>.14</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Volunteers (N = 621)</td>
<td>Self-report of level of efficiency skills</td>
<td>.31</td>
<td></td>
</tr>
<tr>
<td>Schmidt et al., 2006</td>
<td></td>
<td>Volunteers (N = 820)</td>
<td>Self-report of level of medical skills</td>
<td>.65</td>
<td></td>
</tr>
<tr>
<td>Medical skills</td>
<td></td>
<td>Volunteers (N = 820)</td>
<td>Written test measuring knowledge of medical skills (111) items</td>
<td>1.15</td>
<td></td>
</tr>
<tr>
<td>Scherbier, 1997</td>
<td>Groningen</td>
<td>Volunteers (N = 30)</td>
<td>Written test measuring knowledge of medical skills (186) items</td>
<td>1.20</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Volunteers (N = 29)</td>
<td>Written test measuring knowledge of medical skills (132) items</td>
<td>1.89</td>
<td></td>
</tr>
<tr>
<td>Scherbier, 1997</td>
<td></td>
<td>Volunteers (N = 29)</td>
<td>Written test measuring knowledge of medical skills (132) items</td>
<td>1.60</td>
<td></td>
</tr>
<tr>
<td>Remmen et al., 2001</td>
<td>Antwerpen</td>
<td>Volunteers (N = 178)</td>
<td>Written test measuring knowledge of medical skills (132) items</td>
<td>1.26</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Volunteers (N = 262)</td>
<td>Written test measuring knowledge of medical skills (132) items</td>
<td>1.26</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ghent</td>
<td>Volunteers (N = 118)</td>
<td>4 classes: 3rd–7th year</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Volunteers (N = 364)</td>
<td>Written test measuring knowledge of medical skills (132) items</td>
<td>1.60</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Groningen</td>
<td>Volunteers (N = 118)</td>
<td>Written test measuring knowledge of medical skills (132) items</td>
<td>1.26</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Volunteers (N = 115)</td>
<td>Written test measuring knowledge of medical skills (132) items</td>
<td>1.26</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>PBL-school: Volunteers (N = 118)</td>
<td>Written test measuring knowledge of medical skills (186) items</td>
<td>1.20</td>
<td></td>
</tr>
<tr>
<td>Judgments of the quality of medical education</td>
<td></td>
<td>Samples drawn at three occasions: 2004, 2006, and 2008</td>
<td>Rating scale of 30–33 items</td>
<td>.53</td>
<td></td>
</tr>
<tr>
<td>Steenkamp et al., 2004; 2006; 2008</td>
<td>Leiden</td>
<td>PBL-school: Random samples (N = 168)</td>
<td>Rating scale of 30–33 items</td>
<td>.53</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Conventional school: Random samples (N = 184)</td>
<td>Rating scale of 30–33 items</td>
<td>.69</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Utrecht</td>
<td>PBL-school: Random samples (N = 168)</td>
<td>Rating scale of 30–33 items</td>
<td>.75</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rotterdam</td>
<td>Conventional school: Random samples (N = 235)</td>
<td>Rating scale of 30–33 items</td>
<td>.61</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>PBL-school: Random samples (N = 168)</td>
<td>Rating scale of 30–33 items</td>
<td>.61</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Amsterdam UvA</td>
<td>Conventional school: Random samples (N = 197)</td>
<td>Rating scale of 30–33 items</td>
<td>.61</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>PBL-school: Random samples (N = 168)</td>
<td>Rating scale of 30–33 items</td>
<td>.61</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Conventional school: Random samples (N = 160)</td>
<td>Rating scale of 30–33 items</td>
<td>.61</td>
<td></td>
</tr>
</tbody>
</table>
### Summary of Studies Comparing, on Various Outcome Measures, the Performance of Students and Graduates of the Problem-Based Medical Curriculum Under Study and Conventional Medical Schools (Continued)

<table>
<thead>
<tr>
<th>Author</th>
<th>Conventional School</th>
<th>Sample Information (Population/Volunteers/Random Selection, N)</th>
<th>No. of Classes Involved and Level of Training of Participants</th>
<th>Measuring Instrument Information</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amsterdam VU</td>
<td>PBL-school: Random samples (N = 168)</td>
<td>Samples drawn at three occasions: 2004, 2006, and 2008</td>
<td>Rating scale of 30–33 items</td>
<td></td>
<td>.72</td>
</tr>
<tr>
<td>Graduation rates and study duration</td>
<td>Conventional school: Random samples (N = 186)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post et al., 1986</td>
<td>7 conventional schools</td>
<td>PBL-school: Populations</td>
<td>5 classes of graduates from PBL school 1974–1978 and class of 1970 of all other schools</td>
<td>Graduation rates</td>
<td>See text</td>
</tr>
<tr>
<td>Post et al., 1986</td>
<td>7 conventional schools</td>
<td>Conventional schools: Populations</td>
<td>5 classes of graduates from PBL school 1974–1978 and class of 1970 of all other schools</td>
<td>Study duration</td>
<td>See text</td>
</tr>
<tr>
<td>Schmidt et al., 2009</td>
<td>Leiden</td>
<td>PBL-school: Populations (N = 1,331)</td>
<td>9 classes of graduates, entering medical school between 1989 and 1997</td>
<td>Graduation rates</td>
<td>.33</td>
</tr>
<tr>
<td>Amsterdam UvA</td>
<td>Conventional school: Populations (N = 1,333)</td>
<td></td>
<td>9 classes of graduates, entering medical school between 1989 and 1997</td>
<td>Graduation rates</td>
<td>.34</td>
</tr>
<tr>
<td>Amsterdam VU</td>
<td>PBL-school: Populations (N = 1331)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Conventional school: Populations (N = 1,643)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PBL-school: Populations (N = 1,331)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Conventional school: Populations (N = 1,738)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Continued on next page)
Rotterdam  
PBL-school: Populations ($N = 1,331$)  
9 classes of graduates, entering medical school between 1989 and 1997  
Graduation rates  
.35

Utrecht  
Conventional school: Populations ($N = 1,838$)  
PBL-school: Populations ($N = 1,331$)  
9 classes of graduates, entering medical school between 1989 and 1997  
Graduation rates  
.28

Schmidt et al., 2009  
Leiden  
Conventional school: Populations ($N = 1,369$)  
PBL-school: Populations ($N = 1,331$)  
9 classes of graduates, entering medical school between 1989 and 1997  
Study duration  
$- .97^f$

Amsterdam  
UvA  
Conventional school: Populations ($N = 1,333$)  
PBL-school: Populations ($N = 1,331$)  
9 classes of graduates, entering medical school between 1989 and 1997  
Study duration  
$- .68^f$

Amsterdam  
VU  
Conventional school: Populations ($N = 1,643$)  
PBL-school: Populations ($N = 1,331$)  
9 classes of graduates, entering medical school between 1989 and 1997  
Study duration  
$- .64^f$

Rotterdam  
Conventional school: Populations ($N = 1,738$)  
PBL-school: Populations ($N = 1,331$)  
9 classes of graduates, entering medical school between 1989 and 1997  
Study duration  
$- .37^f$

Utrecht  
Conventional school: Populations ($N = 1,838$)  
PBL-school: Populations ($N = 1,331$)  
9 classes of graduates, entering medical school between 1989 and 1997  
Study duration  
$- .78^f$

**Note.** PBL = problem-based learning; MC = multiple-choice.  
aEffect size is expressed as Cohen's $d$. Bold indicates an average $d$, weighted over classes of students and numbers of students.  
bFor a description of the Progress Test, see text.  
cSamples were selected at random from pools of volunteers.  
dThe shorter the processing speed, the better.  
eBelgian medical schools have a 7-year curriculum, so the Ghent and Antwerpen schools contributed 4th- to 7th-year students.  
fThe shorter the study duration, the better.