Adv in Health Sci Educ
DOI 10.1007/s10459-015-9584-2

The role of environmental and individual characteristics in the development of student achievement: a comparison between a traditional and a problem-based-learning curriculum

Stefan K. Schauber · Martin Hecht · Zineb M. Nouns · Adelheid Kuhlmey · Susanne Dettmer

Received: 12 August 2014 / Accepted: 19 January 2015

Abstract In medical education, the effect of the educational environment on student achievement has primarily been investigated in comparisons between traditional and problem-based learning (PBL) curricula. As many of these studies have reached no clear conclusions on the superiority of the PBL approach, the effect of curricular reform on student performance remains an issue. We employed a theoretical framework that integrates antecedents of student achievement from various psychosocial domains to examine how students interact with their curricular environment. In a longitudinal study with N = 1,646 participants, we assessed students in a traditional and a PBL-centered curriculum. The measures administered included students' perception of the learning environment, self-efficacy beliefs, positive study-related affect, social support, indicators of self-regulated learning, and academic achievement assessed through *progress tests*. We compared the relations between these characteristics in the two curricular environments. The results are two-fold. First, substantial relations of various psychosocial domains and their associations with achievement were identified. Second, our analyses indicated that there are no substantial differences between traditional and PBL-based curricula concerning the relational structure of psychosocial variables and achievement. Drawing definite conclusions on the role of curricular-level interventions in the development of

S. K. Schauber (✉)
Institute of Medical Sociology and Rehabilitation Science and Department for Assessment, Charité – Universitätsmedizin Berlin, Luisenstraße 57, 10117 Berlin, Germany
e-mail: stefan.schauber@charite.de

M. Hecht
Institute for Educational Quality Improvement, Humboldt–Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany

Z. M. Nouns
Institute of Medical Education, University of Bern, Konsumstrasse 13, 3010 Bern, Switzerland

A. Kuhlmey · S. Dettmer
Institute of Medical Sociology and Rehabilitation Science, Charité – Universitätsmedizin Berlin, Luisenstraße 57, 10117 Berlin, Germany

Published online: 24 January 2015

student's academic achievement is constrained by the quasi-experimental design as wells as the selection of variables included. However, in the specific context described here, our results may still support the view of student activity as the key ingredient in the acquisition of achievement and performance.

Keywords Problem-based learning · Curricular comparison · Progress test · Structural equation modelling · Achievement · Emotion · Development

Introduction

In medical education, the effect of the educational environment on student achievement has primarily been investigated in comparisons between traditional and problem-based learning curricula. Problem-based learning (PBL), which was introduced in the 1970s (Neufeld and Barrows 1974; Neville and Norman 2007), is centred on learning in small tutorial groups typically consisting of six to ten students (Barrows 1996; Schmidt et al. 2009). These groups meet regularly and "discuss practical or theoretical problems designed by staff" (Moust et al. 2005). The real-world problem scenarios (e.g., in the form of medical reports) presented are a key component of PBL; they provide the focus or starting point for student learning. The goals of PBL are to help students "develop (1) flexible knowledge, (2) effective problem-solving skills, (3) self-directed learning (SDL) skills, (4) effective collaboration skills, and (5) intrinsic motivation" (Hmelo-Silver 2004, p. 235). Different characterisations of PBL are given by Schmidt et al. (2009).

Although PBL has been widely adopted (Schmidt et al. 2009), many studies, meta-analyses and reviews of its effectiveness have found no clear-cut evidence for the superiority of this instructional approach (Albanese and Mitchell 1993; Berkson 1993; Vernon and Blake 1993; Colliver 2000; Newman 2003; Dochy et al. 2003; Hartling et al. 2010; Strobel and van Barneveld 2009; Smits 2002). Rather, there seems to be consensus that PBL has several advantages, but that these advantages are tempered by trade-offs. For instance, relative to students in more traditional learning environments, students in PBL curricula tend to perform better on tasks that focus on clinical skills and applied knowledge, but less well in the domain of basic science knowledge and in standardized tests, such as national licensing examinations (Dolmans and Gijbels 2013; Strobel and van Barneveld 2009). More recently, there has been a noticeable shift from rather "isolated" investigations of "hard" outcome variables (e.g., test performance) to more holistic or systematic considerations of the effects of curricular-level interventions. In this respect, an important observation is that reforms of undergraduate medical education often involve not only a new structure of instruction, but also a new *syllabus* (i.e., *what* is taught). From this perspective, it is hardly surprising that PBL students sometimes perform less well on basic science content (Nouns et al. 2012; but see Drukker et al. 2003) and better in the domain of communicative skills (Koh et al. 2008). Such different profiles of performance may be associated with the differing resources (i.e., time) allocated to certain subjects, contents, or learning objectives (Bergman et al. 2008; Farrow and Norman 2003). Indeed, Norman (2008) has suggested that this change in focus has had consequences even down to the staff involved, who have increasingly been recruited from the social or behavioural sciences. Importantly, as it is difficult to disentangle the effects of PBL from those of change in the *content* taught, the consequences for students in PBL curricula remain an issue (Dolmans and Gijbels 2013; Dolmans et al. 2005).

How do students interact with their educational environment?

Against this background, several authors have raised the concern that research on differences between educational environments often puts too much emphasis on formal academic outcomes. A focus on performances and achievements alone neglects one of the major effects of a curricular intervention, that is, how educational contexts shape students' perceptions, behaviours and personal development (Mattick and Knight 2007; Vanthournout et al. 2009; Nijhuis et al. 2005; van der Veken et al. 2009a, b). Likewise, ten Cate (2001) has argued that most studies have neglected the role of the *interaction* between students and their (educational) environment (e.g., how they engage with their tutors, fellow students, resources, etc.). From this perspective, the finding that different curricula produce, on average, rather comparable outcomes is potentially explained by the observation that students usually make efforts to learn what is required of them, no matter how the curriculum is structured (Norman 2004; ten Cate 2001). As ten Cate (2001) noted: "The primary cause of knowledge acquisition is student activity and not the teaching" (p. 86).

The role of student activity in an instructional context can be viewed as analogous to a well-known issue in doctor-patient relationships. Although a doctor may develop a treatment plan and assess its impact on patients' symptoms, it is the patient herself who takes the pills, sticks to the prescribed doses of medics and routinely goes to physiotherapy. Patients as well as students have a choice to be "adherent" or "compliant" to the selected treatment or not. Likewise, a learning environment based on a PBL approach is beyond doubt an intervention with the *potential* to radically shape students' daily learning activities (Schmidt et al. 2011; but see Struyven et al. 2006). But it is also plausible to expect students to compensate for a less supportive learning environment (ten Cate 2001). For example, regardless of the instructional approach, students will meet in study groups to proactively learn for an exam or rehearse practical skills by familiarizing with auscultation or even taking blood samples from one another—in short, they may themselves generate more "informal" opportunities to learn. Accordingly, there have been increasing efforts to understand students' learning activities beyond the reach of formal curricular interventions (Hommes et al. 2014, 2012). Investigating how learners act within their educational environment may broaden the scientific understanding of why particular outcomes are observed. To this aim, as ten Cate (2001) has suggested, it will likely shift the perspective of research on curricular comparisons from an emphasis on test performance to one of the most important outcomes in education: the *students* themselves with all their motivations, perceptions, efforts, emotions, and so forth.

The idea that the process of learning can be adequately understood only if environmental, social and psychological antecedents are integrated has been a continuing theme in research on self-regulated learning. For instance, there has been substantial interest in research that relates students' learning outcomes to affective, motivational and behavioural characteristics. Recurring themes in this area of research are the relations between academic achievement and students' goals (Diseth 2011; Artino et al. 2012a), perception of the learning environment (Nijhuis et al. 2005; Diseth 2007; Rytkönen et al. 2012), emotions (Artino et al. 2010), motivations (Kusurkar et al. 2013), self-efficacy beliefs (Richardson et al. 2012; Wäschle et al. 2014), and the process of learning (Schmitz and Wiese 2006). Despite the variety of approaches and concept, one main implication of these studies for comparisons between curricula is that, although the contextual factors clearly have an impact on student learning and achievement, learning itself takes place *within* individual persons (Loyens et al. 2008). This means that students set themselves *goals*; they *feel frustrated* when an exam is inappropriately difficult, and they make *efforts* to complete their studies—but they may also *decide* not to. Likewise, dependent on individual characteristics such as prior knowledge or general cognitive abilities, some students benefit from certain forms of instruction while others may not (Kirschner et al. 2006). Put differently, antecedents of achievement are located in the educational setting, in the social environment, and as well as in learners' very own "intra-personal" context (Pekrun 2006; Pekrun et al. 2002; Artino et al. 2012b; Gijbels et al. 2008).

Importantly, also findings from other areas of psychological research highlight the tight connection between affective and cognitive components of human performance (Brose et al. 2012; Burt et al. 1995; Riediger et al. 2011; Sliwinski et al. 2006; Nolen-Hoeksema et al. 2008). These investigations provide ample evidence for the need to take an integrated perspective on the different domains of human functioning in order to gain a better understanding of actual behaviour. Similarly, research on human cognition suggests that motivational and self-regulatory skills (e.g., behavioural control and inhibition) are significantly related to general cognitive abilities (Nisbett et al. 2012; Hofmann et al. 2012). In conclusion, successful learning—like many other human activities—can be regarded as the result of self-regulation (i.e., "goal-directed behaviour, typically within at least a minimum temporal perspective" Hofmann et al. 2012, p. 174), which can in turn be understood as an interplay between key domains of human functioning, such as cognition, emotion and motivation.

What this study adds

Based on this review of the literature we investigate whether similarities in the outcomes of different learning environments go hand-in-hand with equivalent patterns of interrelations in psychological antecedents of achievement. To address this question, we examine the relations between educational outcomes (learning behaviour, study effort, and achievement), emotions (positive study-related affect), appraisals (general self-efficacy and student evaluation of the learning environment) and aspects of the environment (social support) in both, a PBL-focused and a traditional track. This systematization is based on Pekrun (2006) and provides the framework for postulating relationships between variables in this study. More specifically, we focus on the following research questions:

1. Is social support among students positively related to students' perception of the educational environment and control expectancies? (environment → appraisal)
2. Are beneficial appraisals (control expectancies/perception of the environment) related to positive study-related affect? (appraisal → emotion)
3. Is positive study-related affect related to study effort? (emotion → behaviour/activity)
4. Is study effort associated with facets of self-regulated learning and with steeper learning gains across the observation period? (learning behaviour → learning outcomes)
5. Is student collaboration related to perceived social support? (learning behaviour → environment)

Critically, we investigate whether the pattern of interrelations of these variables is dependent on the educational context or not. That is, we investigate the similarities and differences regarding the relational structure between antecedents of academic achievement across cohorts of students from two differently structured educational environments—a PBL-track and a traditional track. In order to test these differential effects, we apply a multiple-group structural equation model in which the groups are the two specific educational contexts under investigation. Furthermore, we designed the study to integrate a large number of students and to follow their development in these two contexts across a considerably long phase of 2 ½ years.

This study thus complements and extends on previous research emphasising both psychosocial facets of (self-regulated) student learning in medical education (Stegers-Jager et al. 2012; Artino et al. 2010; Artino et al. 2012a; Kusurkar et al. 2013) and curricular comparisons—especially with a focus on the effectiveness of PBL (Verhoeven and Verwijnen 1998; van der Veken et al. 2009a, b; Schmidt et al. 2009; Bate et al. 2014). To our knowledge, no previous comparisons of traditional and PBL curricula in medical education research have analysed the relations between the educational context, student characteristics and academic outcomes as outlined above.

Method

Educational context

Since 1999, a reformed medical curriculum (RMC) has been run in parallel to a traditional medical curriculum (TMC) at Charité – Universitätsmedizin Berlin. The RMC can be characterized as a student-centred, vertically integrated curriculum with a strong focus on interpersonal skills (Kiessling et al. 2004; Nouns et al. 2012). Its central instructional approach is PBL, implemented in line with the 5 "ground rules" formulated by Maudsley (1999). The 7-step approach formulated by Schmidt (1983) serves as a basis for the structuring of PBL courses. PBL sessions last about 2 h and are held twice a week. The sessions are facilitated by specially trained tutors from the academic staff with a medical or non-medical background (basic sciences, clinical subjects, and social and behavioural sciences). No student tutors are involved. Due to the legal framework, students in the RMC are not required to take the part of the national licensing examination that focuses mostly on basic science (similar to the United States Medical Licensing Examination Step 1), but obtain a formal equivalence to this exam by the completion of the third academic year. By contrast, students in the TMC have to pass the national licensing examination to enter the "clinical part" of their education. TMC students are expected to take this exam at the end of their second academic year. The TMC can be characterized as teacher-centred, with lectures and seminars being the dominant instructional approach (Kiessling et al. 2004). Sixty students per year are enrolled in the RMC and some 250 students are enrolled twice per year in the TMC. Candidates applying to study medicine at Charité – Universitätsmedizin Berlin can opt to take part in a lottery for a place in the RMC. Further details on the characteristics of the educational context are given by Kiessling et al. (2004), Nouns et al. (2012), and Schauber et al. (2013).

Procedure and administration

We obtained self-report measures and results from progress tests as indicators of academic achievement on five consecutive occasions from April 2011 to April 2013. Progress tests are scheduled in the first week of each semester (April/October) for all medical students at Charité. Test dates are scheduled for groups of about 50–100 students each. Before each test, students were introduced to the present study and handed a one-page self-report questionnaire. This questionnaire tapped basic demographics and self-reports, and included an instruction to generate a unique personal ID. Participants were asked to consent to their progress test data being included in the analyses. Those who consented were asked to transcribe a multi-digit dataset number to the questionnaire. Data acquisition was conducted with the approval and advice of the local data protection authorities. Due to the local directives, we did not acquire further personal data from the students. Consequently, no formal records on study progress (attrition, duration, prior school achievements, etc.) or year of birth are available. The whole procedure was designed to take about 5 min to minimise inference with the general testing procedure. As no patients or physicians were involved and local data protection rules were met, this study was deemed exempt from review by the local ethical review board.

Participants

The present data were obtained from $N = 1{,}646$ students ($N_{TMC} = 1{,}471$ and $N_{RMC} = 175$) in a prospective longitudinal design. Specifically, we gathered data from multiple cohorts of students between the 6th and the 10th semester. This period of training was chosen as both curricula deal primarily with clinical content at this point (i.e., “clinical phase” of the TMC, “2nd part” of the RMC; see Nouns et al. 2012 for details). The form of administration raised several problems: First, technical problems (e.g., loss of questionnaires) led to incomplete data. Second, presumably due to individual errors in transferring the multi-digit dataset number required for data protection purposes, not all of the questionnaires submitted could be combined with test data.

Missing data

Missing data were present at all measurement occasions. To a considerable extent, this could be attributed to factors such as organisational difficulties or transcription and scan errors. We therefore assume that the data are “missing at random” (MAR) and can be adequately handled by full information maximum likelihood (FIML) estimation (Little and Rubin 2002; Enders 2001).

Measures

Environment: enrolment and social support

The main indicator for the curricular context was enrolment to either the reformed, PBL-based, or the traditional medical curriculum. Aside from the curricular context, we adapted two items from the German version of the Interpersonal Support Evaluation List (ISEL; Cohen and Hoberman 1983; Laireiter 1996) to our context. One item targeted emotional support from fellow students (“When I need emotional support, there are other students I can talk to”); the other, informal support/advice (“There is someone I can turn to for advice about handling problems with my studies”).

Appraisal: self-efficacy and students' perception of the learning environment

We administered a shortened version of the general self-efficacy scale by Schwarzer and Jerusalem (1995) as well as a short scale tapping *students' perception of the learning environment.* The latter scale was adapted from Westermann et al. (1996) to the context of medical education (see also Dettmer et al. 2010). Students rated didactics, practical relevance, potential for professional development, and overall course structure and organization on 5-point Likert-type scales.

Emotion: positive study-related affect

Three items forming a short *positive study-related affect* scale were rated on a 5-point Likert-type scale. The items were developed in accordance with Pekrun's framework (Pekrun et al. 2002), which considers positive achievement-related affect to be associated with emotions relating to prospective (e.g., hope), retrospective (e.g., satisfaction) and current achievements (e.g., confidence, enjoyment). The wording was as follows: "I'm confident that I will complete my studies successfully" (prospective), "I enjoy my studies" (current), and "I am satisfied with my studies" (retrospective). These items closely resemble formulations from the positive affect subset of the Achievement Emotion Questionnaire (Pekrun et al. 2011), but were adapted to our context.

Achievement and learning: progress tests and approaches to learning

Two short scales assessed student approaches to learning. First, we implemented a subset of items from the German version of the Motivated Strategies for Learning Questionnaire (MLSQ, Pintrich et al. 1993; Wild and Schiefele 1994) that targeted *metacognitive strategies* (e.g., planning one's efforts) and *collaborative learning* efforts (7 items each subscale). Second, three items tapped student self-reports of s*tudy effort*. These three items were based on the effort-regulation scale (Wild and Schiefele 1994) and reworded for use in our setting of medical education. Specifically, students were asked whether they made deliberate efforts to deal with medical topics and to foster their medical knowledge and skills (e.g., "I use every opportunity to rehearse my clinical skills").

On all occasions, *achievement* was assessed using progress tests (Progress Test Medicine, PTM). Each test contained 200 multiple choice items in single-best answer format which are sampled from a pool of around 5,000 questions covering the entire medical curriculum (including content from both the pre-clinical and clinical phase). Students from all levels of training take the identical test at each measurement occasion. In order to limit recognition effects, all items in a once administered test are withheld for further use for about 2 years. In our analysis students' number-correct scores were used as an indicator of student achievement. Results of several previous studies using data from the PTM suggest that it is a reliable instrument that holds both concurrent and predictive validity (Nouns et al. 2012, 2004 ; Schmidmaier et al. 2010). For details on the Progress Test Medicine, see Nouns and Georg (2010).

Analysis

We approached our research questions in two steps. First, we investigated the measurement properties and stability of the self-report measures obtained across the time span considered. The aim was to evaluate whether variables could be integrated into a single, stable measure in the subsequent structural equation model. This first step was taken to make the subsequent analyses as parsimonious as possible without loss of information. Second, we addressed our main research question, investigating the interdependence structure of the self-report measures and achievement within the two curricular contexts. To that end, the variables were combined in a comprehensive model of antecedents of academic achievement that included measures from the domains of environment, appraisal, emotion and achievement. We analysed the relations between the domains by means of a structural equation model. We describe the rationale of these two steps in more detail in the following paragraphs.

Step 1: measurement properties and stability of measures

The first step in our analysis was carried out instrument-wise (i.e., we assessed properties independently for each variable). In particular, we tested the stability of the variables across the time span considered. The scheme of analysis was identical for all variables. Specifically, we applied a linear mixed-effects model—also called multilevel regression model—(Hox 2002; Gelman and Hill 2007) in which we estimated variance components for three facets: *student* (i.e., stable between-person differences across measurement occasions), *time* (i.e., similar patterns of change over time across students) and *student–time interaction* (differences in patterns of change over time between students). If a variable was stable, we would expect a large share of the variance to be attributable to the student component compared to the other variance components. This relative share of variance was quantified as the *intraclass correlation* (ICC).

Step 2: investigating the interdependence structure of the measures within the two curricula

To address our main research question, we fitted a multiple-group structural equation model to the data. This type of analysis yields *separate* parameters for each group within *a single* comprehensive model. The major advantage over group-specific, independent models is that it is possible to test assumptions on the *equality* of coefficients between groups. In our model, the two groups of interest are the PBL-based reformed medical curriculum (RMC) group and the traditional medical curriculum (TMC) group. For each of these groups, the between-variable relations specified were identical. In a first model, we freely estimated the regression coefficients for each group. This "context-dependence model" accounts for possible differences in the pattern of between-variable relations dependent on the grouping variable (i.e., educational context: RMC/TMC). In a second model, we constrained the magnitude of regression coefficients to be equal across the two groups (RMC/TMC). This "context-independence model" assumes equality of the magnitude of regression coefficients across the two groups. To test for overall group differences in the relation of variables across contexts, we compared the relative fit of the two models to the data (see below for details).

Students' performance on progress tests—our indicator of academic achievement—was expected to improve markedly from semester 6 to semester 10 (Schauber et al. 2013; Verhoeven et al. 2005, 2002). Thus, these scores were specified as a latent growth model (LGM) (Preacher 2008). LGMs can be described as similar to repeated measures analysis of variance (RM-ANOVA), but offer a more flexible approach to modelling change. Technically speaking, like RM-ANOVAs, LGMs make it possible to account for stable between-person differences across the period of observation while estimating a general trend component (i.e., a slope or the effect of "time in education"). Such a trend would, for example, indicate that students generally get "better" over their course of study—for instance, a 5th year student will likely be able to answer more questions on a test on medical knowledge correctly than a 2nd year student. However, RM-ANOVAs assume that these differences between students remain stable across the observed time span. Thus, "high performers" and "low performers" are expected to maintain their relative order throughout the study. Clearly, this is a rather strict assumption, as students may differ not only in their *initial* levels of performance but also in their *gains* in performance. Against this background, LGMs offer clear advantages over RM-ANOVAs as they allow researchers to separate individual differences in initial levels of performance from differences in the "rate" or steepness of growth (e.g., gains or losses) across the period considered. LGMs can thus account for inter-individual (i.e., between-person) differences in intra-individual (i.e., within-person) change. Separating these components is of both theoretical and practical interest—for example, when higher levels of initial performance (e.g., prior knowledge) may be correlated with steeper or more rapid achievement gains, as in the "Matthew effect" ("Whoever has will be given more", Walberg and Tsai 1983). Another example is the investigation of long-term effects of deliberative practice, where continuous efforts may be associated with steeper gains and result in substantive long-term differences in performance (Moulaert et al. 2004; Duvivier et al. 2011; Kulasegaram et al. 2013).

In our analysis, we adopt the view that between-person differences in antecedents of student achievement (i.e., study effort, study-related affect, etc.) are related to between-person differences in achievement *gains*. We thus used the LGM to statistically separate differences between students' achievement *gains* from inter-individual differences in *initial levels* of achievement at the start of our study. Technically speaking, the latter were included as a random effect, allowing for an unbiased estimation of the growth component (i.e., "gains in achievement").

Assessing model fit

In structural equation modelling, several criteria of model quality have to be met. In our study, the models' fit to the data was determined by the Root Mean Square Error of Approximation (RMSEA), the Tucker-Lewis Index (TLI), the Comparative Fit Index (CFI), and the Standardized Root Mean Square Residual (SRMR). The RMSEA compares the model-implied covariances with those observed in the sample. Larger RMSEA values indicate a larger discrepancy between the specified model and the data; values of 0.06 and below are usually considered to indicate good fit. Both the TLI and the CFI test the specified model against an "independence model," which basically assumes the observed variables to be uncorrelated. In general, values close to 0.95 or higher are considered good and values < 0.90 are deemed insufficient. The SRMR is calculated as the standardized differences between the empirical correlations and those predicted by the model. Values of 0.08 and less are considered adequate (for further details, see Hu and Bentler 1999 or Kline 2011).

Data processing and parameter estimation

Pre-processing of data, estimation of measurement properties, and linear mixed-effects modelling (multilevel modelling) was conducted using the R language for statistical computing (R Core Team 2014) and the R package *lme4* (Bates et al. 2014). Structural equation models were estimated in Mplus 6.1 (Muthén and Muthén 1998–2011).

Results

Psychometric properties of variables: reliability within and across measurement occasions

The psychometric properties of the variables used in the analysis are reported in Table 1. The ICCs were large for all obtained variables. In addition, reliability coefficients for most measures seemed to be adequate, with the exception of social support, with Cronbach's alpha = 0.41. All other measures ranged between alpha = 0.66 and alpha = 0.91 across instruments.

Contrasting descriptive statistics for the two curricula

We next averaged each variable (*excluding achievement*) for each person across measurement occasions. We then calculated descriptive statistics for these variables within both curricula (see Table 2). Inspection of these descriptive statistics suggests that, first, although RMC students showed higher achievement levels earlier in the course of study (6th semester: $M_{TMC} = 31.72$ vs. $M_{RMC} = 34.28$), TMC students showed slight benefits

Table 1 Psychometric characteristics of the obtained self-reports

Instrument	Example item	Reliability (Cronbach's alpha within occasion)		ICC
		M	*SD*	
General self-efficacy $N_{(Items)} = 8$	"I can always mange to solve difficult problems if I try hard enough"	0.88	0.02	0.67
Social support $N_{(Items)} = 2$	"When I need emotional support, there are other students I can talk to"	0.41	0.06	0.56
Positive study-related affect $N_{(Items)} = 3$	"I enjoy my studies"	0.66	0.03	0.59
Metacognitive approaches to learning $N_{(Items)} = 7$	"Before studying I consider *what* to learn"	0.73	0.02	0.55
Collaborative learning $N_{(Items)} = 7$	"I ask a fellow student to test me and also ask him or her questions on the content covered"	0.84	0.01	0.65
Study effor $N_{(Items)} = 3$	"I use every opportunity to rehearse my clinical skills"	0.73	0.02	0.70
Students' perception of the learning environment $N_{(Items)} = 4$	"I'm content with the overall course quality"	0.79	0.03	0.71
Achievement/Progress test score	See "Appendix"	0.91	0.02	0.85

Reliability (alpha) for progress tests were calculated within semesters

Alpha Cronbach's alpha, *M* mean, *SD* standard deviation, *ICC* intraclass correlation for student component

Table 2 Descriptive statistics for psychosocial variables and progress tests

	Traditional curriculum (TMC)		Reformed curriculum (RMC)	
	M	*SD*	*M*	*SD*
Self-reports				
Study effort	2.67	0.89	2.87	0.89
Social support	2.82	0.75	2.93	0.69
General self-efficacy	2.74	0.48	2.78	0.53
Positive study-related affect	3.24	0.52	3.31	0.54
Students' perception of the learning environment	2.01	0.54	2.63	0.56
Metacognitive approaches to learning	3.12	0.67	3.12	0.58
Collaborative approaches to learning	2.69	0.83	2.95	0.77
Progress test				
6th semester	31.72	10.42	34.28	10.97
7th semester	36.96	12.24	37.73	12.08
8th semester	42.74	13.43	43.09	12.21
9th semester	47.37	13.88	43.47	12.06
10th semester	51.72	14.80	49.70	12.22

by the 10th semester ($M_{TMC} = 51.72$ vs. $M_{RMC} = 49.70$). However, RMC students reported being more content with their learning environment ($M_{TMC} = 2.01$ vs. $M_{RMC} = 2.63$) and engaging more in collaborative learning ($M_{TMC} = 2.69$ vs. $M_{RMC} = 2.95$).

A multiple-group model examining the role of antecedents of achievement

We next fitted a multiple-group structural equation model. First, we fitted an unrestricted variant of this model that allowed regression coefficients to differ between the two groups (i.e., curricula). Second, we restricted regression coefficients to be equal across these groups. We then inspected the corresponding fit indices. The only fit indices that fell below the suggested values for a good fit—but were still in an acceptable range—were the CFI and the TLI. As the other fit indices indicated good fit, this result can be explained by the fact that a considerable number of relations between variables in our model were expected

Table 3 Fit statistics of the specified models

Model	Number of free parameters	χ^2	df	$\Delta\ \chi^2$ (Δ df)	RMSEA (90 % CI)	CFI	TLI	SRMR
Context-dependence	64	296	116	–	0.043 (0.037 – 0.050)	0.93	0.93	0.07
Context-independence	51	307	129	11 (13) $p = .40$	0.041 (0.036 – 0.048)	0.94	0.93	0.08

df degrees of freedom; *RMSEA* Root Mean Square Error of Approximation, *TLI* Tucker-Lewis Index, *CFI* Comparative Fit Index, *SRMR* Standardized Root Mean Square Residual to be (near to) zero. Consequently, testing against a null model, as done in the CFI and TFI, will—by definition—produce lower indices than in a situation where correlations between all variables are expected, which is not the case for the present theoretical framework. Thus, we concluded that both models had an adequate fit to the data. Importantly, there was no significant difference in the fit statistics of the two models (see Table 3). Hence, we accepted the more parsimonious model, which assumed equality between the curricula. This result indicates that any differences in the pattern of relations between variables associated with educational outcomes in the PBL-oriented versus the traditional learning environment were statistically negligible.

The unstandardized results of the constrained model presented in Fig. 1 are, generally speaking, in line with the hypotheses derived from Pekrun's framework. In the following, we report standardized regression coefficients. In particular, social support was related to general self-efficacy (RMC: $\beta_{std} = 0.22$, S.E. = 0.03; TMC: $\beta_{std} = 0.24$, S.E. = 0.03) and student perception of the learning environment (RMC: $\beta_{std} = 0.30$, S.E. = 0.03; TMC: $\beta_{std} = 0.30$, S.E. = 0.03). Both general self-efficacy (RMC: $\beta_{std} = 0.27$, S.E. = 0.03; TMC: $\beta_{std} = 0.26$, S.E. = 0.02) and student perception of the learning environment (RMC: $\beta_{std} = 0.35$, S.E. = 0.03; TMC: $\beta_{std} = 0.35$, S.E. = 0.02) were positively related to positive study-related affect. In turn, positive study-related affect was positively related to self-reported study effort (RMC: $\beta_{std} = 0.24$, S.E. = 0.03; TMC: $\beta_{std} = 0.26$, S.E. = 0.03), which encompasses positive approaches to studying. Study effort was positively related to achievement gains across the course of study (RMC: $\beta_{std} = 0.44$, S.E. = 0.12; TMC: $\beta_{std} = 0.30$, S.E. = 0.05). In addition, collaborative learning was positively related to perceived social support (RMC: $\beta_{std} = 0.35$, S.E. = 0.04; TMC: $\beta_{std} = 0.35$, S.E. = 0.02). When the regression coefficients were fixed to be equal across contexts, the only difference in standardized coefficients was a stronger relation between self-reported study effort and achievement gains ($\beta_{(TMC)std} = 0.30$ vs. $\beta_{(RMC)std} = 0.45$).

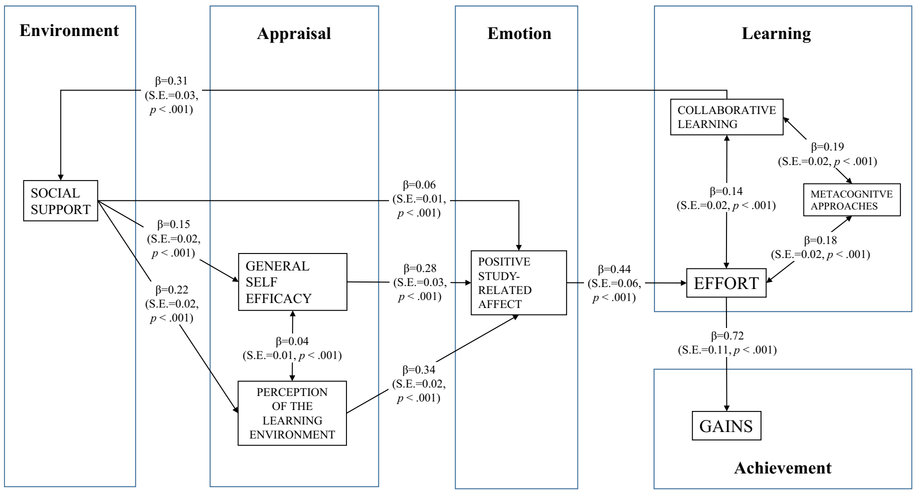

Fig. 1 Unstandardized results for the independence model. *GAINS* signify intra-individual development of performances (i.e., achievement gains). For standardized regression coefficients refer to the text

In summary, these results support our hypotheses on the interrelation between environmental and student characteristics across the course of study. In addition, a multiple-group structural equation model assuming identity across contexts in the relations between the measured characteristics was best able to explain the data.

Discussion

Our study examined the interplay between environmental and individual characteristics in the development of student achievement. Previous research on antecedents of academic achievement (Pekrun 2006) served as a theoretical basis for postulating relationships between the learning environment, individual appraisal, emotion, and achievement. Importantly, we tracked a large number of students ($N = 1{,}646$) and their development across a phase of 2 ½ years of medical education. We further investigated the effect of two curricula (problem-based vs. traditional) on both the outcomes (means) and the interrelations of variables (structural equation model). A multiple-group structural equation model that assumed equivalence of relations between antecedents of achievement and outcome measures fitted the data appropriately. Hence, we found no evidence that the interplay between students' emotions, evaluations, perceptions on the one hand and outcomes on the other hand is dependent on the educational context.

In terms of means alone, our results replicate previous findings and confirm that PBL curricula tend to be associated with more favourable student perceptions of the learning environment (Norman and Schmidt 2000; Kaufman and Mann 1996). The results of comparisons based solely on achievement-related data likewise replicate previous findings (Newman 2003). Specifically, there seems to be a rather small difference in academic achievement in favour of TMC students. However, we found no differences in levels of study-related positive affect or study effort (see also Cohen-Schotanus et al. 2008). One of the most valuable finding is that in both contexts gains in achievement is related to self-reported study effort which, in turn, is related to beneficial approaches to learning such as collaborative studying or metacognitive approaches to learning. Further, our results suggest that positive-study related affect plays a central role in adopting such favourable learning activities. In addition, our results support the idea that engaging in collaborative studying seems not only to have benefits in terms of learning but also shows associations with perceived social support among students, which again is related to appraisals such as higher self-efficacy believes and a more positive perception of the learning environment. It should be pointed out that most of the variables we considered remained remarkably stable across the course of study. Although this finding may be expected for measures such as general self-efficacy, the "trait-like" character of other variables (such as positive study-related affect and the measures of approaches to learning) has— to our knowledge—not been addressed in previous research.

Our findings contribute to the understanding of the similarities in medical students' learning gains within different curricular contexts in several respects. First, our results strengthen the view that self-regulatory processes (including affective, cognitive and behavioural factors) play a crucial role in medical students' acquisition of knowledge (Artino et al. 2012b). Second, our results illustrate—at least for the time-span covered here—that certain aspects of self-regulatory processes (e.g., positive study-related affect; collaborative learning) are not necessary linked to the curricular context. Third, shifting the perspective from outcome measures alone to more integrated considerations including psychological variables implies that the effect of curricular intervention on student achievement may also depend on characteristics of the group of students typically recruited.

Against this background, it is important to note that medical students tend to be high performers who usually are expected to have high levels of cognitive resources—which have, in turn, been linked to good self-regulatory skills. Thus, interventions aimed at improving such skills may only have substantial effects for students with lower levels of these abilities and resources. In turn, this may allow for the speculation that as long as the educational context satisfies certain needs (e.g., valid feedback, reasonable consequences for success and failure, support of autonomy) students with higher self-regulatory abilities will make use of them regardless of the curricular environment. From this perspective, it is interesting to note that TMC students reported lower satisfaction ratings, but not lower ratings of positive study-related affect. In other words, they seem—to some degree—to separate their evaluation of teaching quality from the enjoyment they derive from dealing with the subject of medicine. In sum, although the curricular environment (RMC vs. TMC) was associated with considerable differences in self-reported course-related contentedness, there were no corresponding differences in the other variables.

Limitations and strengths

Despite these critical contributions to the literature, it has to be acknowledged that our results are based on quasi-experimental data obtained in a particular cultural context. Thus, our capacity to draw conclusions on the underlying mechanisms is limited, as is the potential for generalization to other contexts. Moreover, several limitations in the instruments administered warrant consideration. For instance, given the limited time available to administer our questionnaire, it was not feasible to use more detailed instruments. This is a clear deficit as more fine-grained instruments would offer more detailed insights into the differential effects of achievement emotions (Artino et al. 2012b). The indicators of academic outcomes we obtained may also be limited, as the usefulness of formative tests in the assessment of educational outcomes has been a matter of concern (ten Cate 2001). Likewise, in a recent study, Stegers-Jager et al. (2012) used self-reports of course attendance as an indicator of study effort. Their results suggest that such variables might be a more feasible and accurate indicator of study effort than the adapted self-reports used here.

However, the pattern of results observed in our study is largely in line with previous research on the associations between affective and behavioural components and achievement. Indeed, most of the relationships observed were of moderate magnitude (i.e., correlations in the range of $.3 > r < .4$, corresponding to a Cohen's d of approximately $.63 > d < 0.87$, Borenstein 2009; Wolf 1986). In the context of our study, this might not be a surprising finding, as most of the scales administered were relatively short and their lower reliability, especially of the social-support scale (alpha = 0.42), limits the magnitude of the effects detectable. Against this background, the magnitude of relations observed may be deemed in a satisfactory range and thus interpreted as a validation of our approach.

Another limitation relates to the assessment of outcomes (i.e., academic achievement) by means of progress tests. Proponents of PBL frequently emphasize its benefits for students' flexible application of knowledge (Schmidt et al. 2009; Hmelo-Silver 2004). Yet progress tests may not be able to detect this facet of learning outcomes. Furthermore, although medical knowledge is commonly assumed to form the basis for more complex skills and competencies (e.g., clinical reasoning, history taking, physical examination, etc.), it is not clear in how far the results reported here can be generalized to other domains of student learning. In addition, it was only possible to include a selection of antecedents of academic achievement (4 items focusing on learning environment; 2 on social support) that have been discussed in the literature. Other important contextual factors that were not included in our analyses but have been shown to influence student's approaches to learning include perceptions of the workload and the assessment system, as well as the clarity of goals (Beaten et al. 2010). Hence, the finding of no differences in outcomes might indeed be explained by other variables than those investigated in our model.

Our study focused on a relatively late phase of undergraduate medical education. By this time, the student cohort may already be highly selected, with some students already having dropped out, potentially masking some effects (Schmidt et al. 2009). Complementing this study by research that focuses on an earlier phase of medical education—the transition to medical school—may shed light on crucial aspects of students' development throughout their medical education. To this aim, taking a more fine-grained or micro-analytical approach to investigate students' development in psychosocial domains and their relation to the acquisition of competence would be a promising area of future research (Vanthournout et al. 2009; Dolmans and Gijbels 2013). Such a perspective may enhance our understanding of how the structure of medical education is related to changes in students' thinking, acting, and feeling.

Conclusion

Our results suggest that an approach integrating both psychological variables from various domains and environmental factors is needed to understand (and influence) the learning and achievement of medical students. Although this approach may introduce some complexity, many healthcare professionals are already familiar with analogical conceptualisations which focus on the interplay of biological, environmental, and psychosocial antecedents of health and disease—for example, when dealing with bio-psychosocial models of disease (Engel 1977).

In conclusion, the results of this study are in line with earlier speculations that large-scale interventions, such as curricular reform, that aim to alter students' learning activities may not necessarily have the intended benefits over more traditional learning environments. Rather, students make substantial efforts to reach their goals and succeed in their studies anyway—at least in the context described here. Of course, this does not mean that any change or reform is inevitably unnecessary (e.g., cf. van der Veken et al. 2009a, b). Nevertheless, in our opinion, these findings highlight the need to focus on questions such as *why, which,* and *when* a specific *content* needs to be acquired by medical students—and to allocate resources accordingly. On this basis, instructional approaches can be employed that facilitate the acquisition of specific knowledge, skills and competencies both appropriately and efficiently for the specific group of students that enters medical education.

Acknowledgments The work of Stefan K. Schauber was funded by the German Federal Ministry of Education and Research (BMBF) within the project "Competence Acquisition and Learning Trajectories in Medical Training" (Grant 01JG1055).

Appendix: Example Items

Example item from the clinical knowledge subset

Three hours after carrying a 130 kg washing machine, a 24-year old patient (180 cm, 60 kg) feels an intense pain in the right side of his chest. The initial pain abates after only a few minutes but a light feeling of pressure in the right half of the thorax develops, with pain associated with breathing that increases over time. Eventually an emergency doctor is notified. He finds slight shortness of breath, reduced breathing sounds and percussive response on the right side, a heart rate of 125/min, blood pressure of 120/80 mmHg and a blood oxygen saturation level of 95 % (under normal conditions).

What working diagnosis do you suggest?

a) lung embolism
b) heart attack
c) pneumonia
d) pleural effusion
e) pneumothorax

Organ system: the respiratory tract
Medical discipline: anaesthesiology, emergency and intensive care medicine
Correct answer: *e*

Example item from the biomedical knowledge subset

Which of the following hormones is integral in compensating for a considerable drop in blood volume?

a) adrenaline
b) aldosterone
c) thyroxine
d) nitric oxide

Organ system: hormones, metabolism
Medical discipline: physiology, physics
Correct answer: *b*

References

Albanese, M. A., & Mitchell, S. (1993). Problem-based learning: A review of literature on its outcomes and implementation issues. *Academic Medicine, 68*(1), 52–81.

Artino, A. R., Dong, T., DeZee, K. J., Gilliland, W. R., Waechter, D. M., Cruess, D., & Durning, S. J. (2012a). Achievement goal structures and self-regulated learning: relationships and changes in medical school. *Academic Medicine, 87*(10), 1375–1381.

Artino, A. R., Holmboe, E. S., & Durning, S. J. (2012b). Control-value theory: Using achievement emotions to improve understanding of motivation, learning, and performance in medical education: AMEE Guide No. 64. *Medical Teacher, 34*(3), e148–e160.

Artino, A. R., La Rochelle, J. S., & Durning, S. J. (2010). Second-year medical students' motivational beliefs, emotions, and achievement. *Medical Education, 44*(12), 1203–1212.

Baeten, M., Kyndt, E., Struyven, K., & Dochy, F. (2010). Using student-centred learning environments to stimulate deep approaches to learning: Factors encouraging or discouraging their effectiveness. *Educational Research Review, 5*(3), 243–260.

Barrows, H. S. (1996). Problem-based learning in medicine and beyond: A brief overview. *New Directions for Teaching and Learning, 1996*(68), 3–12.

Bate, E., Hommes, J., Duvivier, R., & Taylor, D. C. M. (2014). Problem-based learning (PBL): Getting the most out of your students—Their roles and responsibilities: AMEE Guide No. 84. *Medical Teacher, 36*(1), 1–12.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4*, Version 1.0-5. http://CRAN.R-project.org/package=lme4. Accessed 22 Jan 2015.

Bergman, E. M., Prince, K. J., Drukker, J., van der Vleuten, C. P. M., & Scherpbier, A. J. (2008). How much anatomy is enough? *Anatomical Sciences Education, 1*(4), 184–188.

Berkson, L. (1993). Problem-based learning: Have the expectations been met? *Academic Medicine, 68*(10), S79.

Borenstein, M. (2009). *Introduction to meta-analysis*. Chichester: Wiley.

Brose, A., Schmiedek, F., Lövdén, M., & Lindenberger, U. (2012). Daily variability in working memory is coupled with negative affect: The role of attention and motivation. *Emotion, 12*(3), 605.

Burt, D. B., Zembar, M. J., & Niederehe, G. (1995). Depression and memory impairment: a meta-analysis of the association, its pattern, and specificity. *Psychological Bulletin, 117*(2), 285.

Cohen, S., & Hoberman, H. M. (1983). Positive events and social supports as buffers of life change stress. *Journal of Applied Social Psychology, 13*(2), 99–125.

Cohen-Schotanus, J., Muijtjens, A. M. M., Schönrock-Adema, J., Geertsma, J., & van der Vleuten, C. P. M. (2008). Effects of conventional and problem-based learning on clinical and general competencies and career development. *Medical Education, 42*(3), 256–265.

Colliver, J. A. (2000). Effectiveness of problem-based learning curricula: research and theory. *Academic Medicine, 75*(3), 259–266.

Dettmer, S., Kuhlmey, A., & Schulz, S. (2010). Karriere- und Lebensplanung: Gehen oder bleiben? *Dtsch Arztebl International, 107*(1-2), A-30.

Diseth, Å. (2007). Approaches to learning, course experience and examination grade among undergraduate psychology students: testing of mediator effects and construct validity. *Studies in Higher Education, 32*(3), 373–388.

Diseth, Å. (2011). Self-efficacy, goal orientations and learning strategies as mediators between preceding and subsequent academic achievement. *Learning and Individual Differences, 21*(2), 191–195.

Dochy, F., Segers, M., Van den Bossche, P., & Gijbels, D. (2003). Effects of problem-based learning: A meta-analysis. *Learning and Instruction, 13*(5), 533–568.

Dolmans, D. H. J. M., de Grave, W., Wolfhagen, I. H. A. P., & van der Vleuten, C. P. M. (2005). Problem-based learning: future challenges for educational practice and research. *Medical Education, 39*(7), 732–741.

Dolmans, D., & Gijbels, D. (2013). Research on problem-based learning: Future challenges. *Medical Education, 47*(2), 214–218.

Drukker, J., Hylkema, N., Prince, K. J. A. H., Scherpbier, A. J. J. A., Van der Vleuten, C. P., & van Mameren, H. (2003). Does problem-based learning lead to deficiencies in basic science knowledge? An empirical case on anatomy. *Medical Education, 37*(1), 15–21.

Duvivier, R. J., van Dalen, J., Muijtjens, A. M., Moulaert, V. R., van der Vleuten, C. P. M., & Scherpbier, A. J. (2011). The role of deliberate practice in the acquisition of clinical skills. *BMC Medical Education, 11*(1), 101.

Enders, C. K. (2001). A primer on maximum likelihood algorithms available for use with missing data. *Structural Equation Modeling, 8*(1), 128–141.

Engel, G. (1977). The need for a new medical model: a challenge for biomedicine. *Science, 196*(4286), 129–136.

Farrow, R., & Norman, G. (2003). The effectiveness of PBL: the debate continues. Is meta-analysis helpful? *Medical Education, 37*(12), 1131–1132.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models. Analytical methods for social research*. Cambridge, New York: Cambridge University Press.

Gijbels, D., Segers, M., & Struyf, E. (2008). Constructivist learning environments and the (im)possibility to change students' perceptions of assessment demands and approaches to learning. *Instructional Science, 36*, 431–443.

Hartling, L., Spooner, C., Tjosvold, L., & Oswald, A. (2010). Problem-based learning in pre-clinical medical education: 22 years of outcome research. *Medical Teacher, 32*(1), 28–35.

Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review, 16*(3), 235–266.

Hofmann, W., Schmeichel, B. J., & Baddeley, A. D. (2012). Executive functions and self-regulation. *Trends in Cognitive Sciences, 16*(3), 174–180.

Hommes, J., Bossche, P., Grave, W., Bos, G., Schuwirth, L., & Scherpbier, A. (2014). Understanding the effects of time on collaborative learning processes in problem based learning: a mixed methods study. *Advances in Health Sciences Education, 19*(4), 541–563.

Hommes, J., Rienties, B., Grave, W., Bos, G., Schuwirth, L., & Scherpbier, A. (2012). Visualising the invisible: a network approach to reveal the informal social side of student learning. *Advances in Health Sciences Education, 17*(5), 743–757.

Hox, J. J. (2002). *Multilevel analysis: Techniques and applications. Quantitative methodology series*. Mahwah, NJ: Lawrence Erlbaum Associates.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55.

Kaufman, D. M., & Mann, K. V. (1996). Comparing students' attitudes in problem-based and conventional curricula. *Academic Medicine, 71*(10), 1096–1099.

Kiessling, C., Schubert, B., Scheffner, D., & Burger, W. (2004). First year medical students' perceptions of stress and support: a comparison between reformed and traditional track curricula. *Medical Education, 38*(5), 504–509.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75–86.

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York: Guilford Press.

Koh, G. C.-H., Khoo, H. E., Wong, M. L., & Koh, D. (2008). The effects of problem-based learning during medical school on physician competency: a systematic review. *Canadian Medical Association Journal, 178*(1), 34–41.

Kulasegaram, K. M., Grierson, L. E. M., & Norman, G. R. (2013). The roles of deliberate practice and innate ability in developing expertise: evidence and implications. *Medical Education, 47*(10), 979–989.

Kusurkar, R. A., ten Cate, T. J., Vos, C. M. P., Westers, P., & Croiset, G. (2013). How motivation affects academic performance: a structural equation modelling analysis. *Advances in Health Sciences Education, 18*(1), 57–69.

Laireiter, A.-R. (1996). *Skalen Soziale Unterstützung.: SSU*. Mödling: Dr. Schuhfried.

Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data* (2nd ed). *Wiley series in probability and statistics*. Hoboken, N.J: Wiley.

Loyens, S. M. M., Magda, J., & Rikers, R. M. J. P. (2008). Self-directed learning in problem-based learning and its relationships with self-regulated learning. *Educational Psychology Review, 20*(4), 411–427.

Mattick, K., & Knight, L. (2007). High-quality learning: Harder to achieve than we think? *Medical Education, 41*(7), 638–644.

Maudsley, G. (1999). Do we all mean the same thing by "problem-based learning"? A review of the concepts and a formulation of the ground rules. *Academic Medicine, 74*(2), 178–185.

Moulaert, V., Verwijnen, M. G. M., Rikers, R., & Scherpbier, A. J. (2004). The effects of deliberate practice in undergraduate medical education. *Medical Education, 38*(10), 1044–1052.

Moust, J. H. C., van Berkel, H. J. M., & Schmidt, H. G. (2005). Signs of erosion: Reflections on three decades of problem-based learning at Maastricht University. *Higher Education, 50*(4), 665–683.

Muthén, L., & Muthén, B. (1998–2011). *Mplus User's Guide* (6th ed.). Los Angeles, CA: Muthén & Muthén.

Neufeld, V. R., & Barrows, H. S. (1974). The "McMaster Philosophy": an approach to medical education. *Journal of Medical Education, 49*(11), 1040–1050.

Neville, A. J., & Norman, G. R. (2007). PBL in the undergraduate MD Program at McMaster University: Three iterations in three decades. *Academic Medicine, 82*(4), 370–374.

Newman, M. (2003). A pilot systematic review and meta-analysis on the effectiveness of problem-based learning. *On behalf of the Campbell Collaboration Systematic Review Group on the Effectiveness of Problem-based Learning. Newcastle upon Tyne, UK: Learning and Teaching Support Network-01, University of Newcastle.*

Nijhuis, J. H., Segers, M. R., & Gijselaers, W. (2005). Influence of redesigning a learning environment on student perceptions and learning strategies. *Learning Environments Research, 8*(1), 67–93.

Nisbett, R. E., Aronson, J., Blair, C., Dickens, W., Flynn, J., Halpern, D. F., & Turkheimer, E. (2012). Intelligence: New findings and theoretical developments. *American Psychologist, 67*(2), 130–159.

Nolen-Hoeksema, S., Wisco, B. E., & Lyubomirsky, S. (2008). Rethinking rumination. *Perspectives on Psychological Science, 3*(5), 400–424.

Norman, G. (2004). Editorial: Beyond PBL. *Advances in Health Sciences Education, 9*(4), 257–260.

Norman, G. (2008). Problem-based learning makes a difference. But why? *Canadian Medical Association Journal, 178*(1), 61–62.

Norman, G. R., & Schmidt, H. G. (2000). Effectiveness of problem-based learning curricula: theory, practice and paper darts. *Medical Education, 34*(9), 721–728.

Nouns, Z. M., & Georg, W. (2010). Progress testing in German speaking countries. *Medical Teacher, 32*(6), 467–470.

Nouns, Z., Hanfler, S., Brauns, K., Foeller, T., Fuhrmann, S., Koelbel, S., & Mertens, A. (2004). Do progress tests predict the outcome of national exams? *AMEE Conference 2004 Edinburgh. Short Communication 2F3*.

Nouns, Z., Schauber, S., Witt, C., Kingreen, H., & Schüttpelz-Brauns, K. (2012). Development of knowledge in basic sciences: a comparison of two medical curricula. *Medical Education, 46*(12), 1206–1214.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*(4), 315–341.

Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The Achievement Emotions Questionnaire (AEQ). *Contemporary Educational Psychology, 36*(1), 36–48.

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist, 37*(2), 91–105.

Pintrich, P. R., Smith, D. A. F., García, T., & McKeachie, W. J. (1993). Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educational and Psychological Measurement, 53*(3), 801–813.

Preacher, K. J. (2008). *Latent growth curve modeling*. Thousand Oaks: Sage.

R Core Team. (2013). *R: A Language and Environment for Statistical Computing*. Vienna, Austria. http://www.R-project.org/

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin, 138*(2), 353–387. doi:10.1037/a0026838.

Riediger, M., Wrzus, C., Schmiedek, F., Wagner, G. G., & Lindenberger, U. (2011). Is seeking bad mood cognitively demanding? Contra-hedonic orientation and working-memory capacity in everyday life. *Emotion, 11*(3), 656.

Rytkönen, H., Parpala, A., Lindblom-Ylänne, S., Virtanen, V., & Postareff, L. (2012). Factors affecting bioscience students' academic achievement. *Instructional Science, 40*(2), 241–256.

Schauber, S. K., Hecht, M., Nouns, Z. M., & Dettmer, S. (2013). On the role of biomedical knowledge in the acquisition of clinical knowledge. *Medical Education, 47*(12), 1223–1235.

Schmidmaier, R., Holzer, M., Angstwurm, M., Nouns, Z., Reincke, M., & Fischer, M. R. (2010). Using the Progress Test Medizin (PTM) for evaluation of the Medical Curriculum Munich (MeCuM). *GMS Zeitschrift für medizinische Ausbildung*, *27*(5), 1–14.

Schmidt, H. G. (1983). Problem-based learning: Rationale and description. *Medical Education, 17*(1), 11–16.

Schmidt, H. G., Rotgans, J. I., & Yew, E. H. J. (2011). The process of problem-based learning: what works and why. *Medical Education, 45*(8), 792–806.

Schmidt, H. G., van der Moldem, H. T., te Winkel, W. W., & Wijen, W. H. (2009). Constructivist, problem-based learning does work: A meta-analysis of curricular comparisons involving a single medical school. *Educational Psychologist, 44*(4), 227–249.

Schmitz, B., & Wiese, B. S. (2006). New perspectives for the evaluation of training sessions in self-regulated learning: Time-series analyses of diary data. *Contemporary Educational Psychology, 31*(1), 64–96.

Schwarzer, R., & Jerusalem, M. (1995). Generalized Self-Efficacy scale. In J. Weinman, S. Wright, & M. Johnston (Eds.), *Measures in health psychology: A user's portfolio. Causal and control beliefs* (pp. 35–37). Windsor: NFER-NELSON.

Sliwinski, M. J., Smyth, J. M., Hofer, S. M., & Stawski, R. S. (2006). Intraindividual coupling of daily stress and cognition. *Psychology and Aging, 21*(3), 545.

Smits, P. B. A. (2002). Problem based learning in continuing medical education: a review of controlled evaluation studies. *BMJ, 324*(7330), 153–156.

Stegers-Jager, K. M., Cohen-Schotanus, J., & Themmen, A. P. N. (2012). Motivation, learning strategies, participation and medical school performance. *Medical Education, 46*(7), 678–688.

Strobel, J., & van Barneveld, A. (2009). When is PBL more effective? A Meta-synthesis of meta-analyses comparing PBL to conventional classrooms. *Interdisciplinary Journal of Problem-based Learning, 3*(1), 45–58.

Struyven, K., Dochy, F., Janssens, S., & Gielen, S. (2006). On the dynamics of students' approaches to learning: The effects of the teaching/learning environment. *Learning and Instruction, 16*, 279–294.

ten Cate, O. (2001). What happens to the student? The neglected variable in educational outcome research. *Advances in Health Sciences Education, 6*(1), 81–88.

van der Veken, J., Valcke, M., de Maeseneer, J., & Derese, A. (2009a). Impact of the transition from a conventional to an integrated contextual medical curriculum on students' learning patterns: A longitudinal study. *Medical Teacher, 31*(5), 433–441.

van der Veken, Jos, Valcke, M., de Maeseneer, J., Schuwirth, L., & Derese, A. (2009b). Impact on knowledge acquisition of the transition from a conventional to an integrated contextual medical curriculum. *Medical Education, 43*(7), 704–713.

Vanthournout, G., Donche, V., Gijbels, D., & Petegem, P. (2009). Alternative data-analysis techniques in research on student learning: illustrations of a person-oriented and developmental perspectives. *Reflecting Education, 5*(2), 35–51.

Verhoeven, B. H., Snellen-Balendong, H. A., Hay, I. T., Boon, J. M., Van Der Linde, M. J., Blitz-Lindeque, J. J., & Scherpbier, A. (2005). The versatility of progress testing assessed in an international context: a start for benchmarking global standardization? *Medical Teacher, 27*(6), 514–520.

Verhoeven, B. H., & Verwijnen, G. M. (1998). An analysis of progress test results of PBL and non-PBL students. *Medical Teacher, 20*(4), 310–316.

Verhoeven, B. H., Verwijnen, G. M., Scherpbier, A., & Van der Vleuten, C. P. (2002). Growth of medical knowledge. *Medical Education, 36*(8), 711–717.

Vernon, D. T., & Blake, R. L. (1993). Does problem-based learning work? A meta-analysis of evaluative research. *Academic Medicine, 68*(7), 550–563.

Walberg, H. J., & Tsai, S.-L. (1983). Matthew effects in education. *American Educational Research Journal, 20*(3), 359–373.

Wäschle, K., Allgaier, A., Lachner, A., Fink, S., & Nückles, M. (2014). Procrastination and self-efficacy: Tracing vicious and virtuous circles in self-regulated learning. *Learning and Instruction, 29*, 103–114.

Westermann, R., Heise, E., Spies, K., & Trautwein, U. (1996). Identifikation und Erfassung von Komponenten der Studienzufriedenheit. [Identifying and assessing components of student satisfaction.]. *Psychologie in Erziehung und Unterricht, 43*(1), 1–22.

Wild, K.-P., & Schiefele, U. (1994). Lernstrategien im Studium: Ergebnisse zur Faktorenstruktur und Reliabilität eines neuen Fragebogens. *Zeitschrift für Differentielle und Diagnostische Psychologie, 15*, 185–200.

Wolf, F. M. (1986). *Meta-Analysis: Quantitative Methods for Research Synthesis. Meta-analysis: Quantitative Methods for Research Synthesis*. Thousand Oaks: SAGE Publications.