Doctoral thesis submitted to
the Faculty of Behavioural and Cultural Studies
Heidelberg University
in partial fulfillment of the requirements of the degree of
Doctor of Philosophy (Dr. phil.)
in Psychology

Title of the publication-based thesis

Supply and Use of Evidence-Based Learning Activities to Improve Teaching and Learning at the University Level

presented by
M.Sc.-Psych. Eva Sophie Bosch

year of submission
2024

Dean: Prof. Dr. Guido Sprenger
Advisor: Prof. Dr. Birgit Spinath

Acknowledgments

First and foremost, I sincerely thank my advisor, Prof. Dr. Birgit Spinath, for her valuable feedback, her constant interest in my research, her confidence in my work, and the freedom she gave me to work on my doctorate in my own way.

I thank Prof. Dr. Oliver Dickhäuser for agreeing to act as a referee for my dissertation and for his encouraging feedback on several of my scientific presentations, especially the first one.

I would like to express my gratitude to all my colleagues for fostering a collaborative working environment, for engaging in constructive discussions, and for enriching our office time. I especially thank my friend and former colleague, Dr. Cordelia Menz, who supported me during the initial stages of my scientific work and agreed to provide feedback on this dissertation. Thank you! I thank Dr. Eva Seifried for her feedback and collaboration on one of the articles included in my dissertation. I further thank Dr. Jelena-Sophie Siebert for her feedback on presentations as well as for her enthusiastic openness to every discussion. Additionally, I thank Dr. Jane Zagorski for her proofreading.

I would like to thank all my Heidelberg friends, especially Jule and Stefan, Julia K., Julia S., Nora, and Vera for enjoying life with me, engaging in conversations about almost everything, participating in countless joint activities, and offering their friendship and support. A special and heartfelt thanks goes to Sophie-Lina Fiebig: Thank you for consistently providing a listening ear, your companionship, interest, ideas, and care.

Finally, I thank my partner Florian for his patience, love, and support as well as my family for always believing in me.

Summary

The aim of this dissertation was to contribute to the understanding of university students' self-regulated learning and the effectiveness of specific learning activities in a real-world learning setting. In university settings in particular, learning is not only simply delivered by a teacher and absorbed by students but learning success largely depends on students' behavior. Hence, it is necessary to analyze what students do in a real-world learning setting and how these activities are related to successful learning. To do so, I assessed students' learning behavior and motivation while attending a lecture class and their association with learning outcomes. In this dissertation, several evidence-based learning activities were implemented in a large university course and their use as well as learning outcomes were evaluated empirically across five cohorts of students between spring 2018 and spring 2022.

The dissertation builds on a supply-use model of learning in higher education. In addressing all parts of the model, I describe learning in a university setting and models of self-regulated learning that fit into this context while also discussing different desirable learning outcomes. I present findings on the role of evidence-based learning activities as well as students' individual prerequisites for academic success. Finally, I present my own empirical findings and, in the last section, I discuss how they can be placed in the context of the state of research. In all three peer-reviewed and internationally published Papers, the use of specific evidence-based learning activities was assessed in large university lecture classes. I continuously aimed to improve the understanding of learning activity use by assessing students' learning intentions (Paper 1), adapting the assessment of activity use (Papers 2 and 3), and identifying different activity use patterns (Paper 3). Students who used many learning activities gained more knowledge beyond motivation and prior achievement (Papers 1 and 3), they further experienced less motivational decline over the course of one semester (Paper 2), and acquired a more accurate self-assessment of their own knowledge (Paper 3).

List of Papers Included in the Publication-Based Dissertation

Paper 1

Bosch, E., Seifried, E., & Spinath, B. (2021). What successful students do: Evidence-based learning activities matter for students' performance in higher education beyond prior knowledge, motivation, and prior achievement. *Learning and Individual Differences*, *91*, 102056. https://doi.org/10.1016/j.lindif.2021.102056

Paper 2

Bosch, E., & Spinath, B. (2023a). Students' motivation in an online and a face-to-face semester: A comparison of initial level, development and use of learning activities. *Zeitschrift für Psychologie, 231*(2), 93–102. https://doi.org/10.1027/2151-2604/a000519.

Paper 3

Bosch, E., & Spinath, B. (2023b). What evidence-based learning activities help students acquire knowledge, correct confidence in their own knowledge, and accurate self-assessment? *Learning and Individual Differences, 108,* 102374. https://doi.org/10.1016/j.lindif.2023.102374

Contents

1. Introduction

Higher education is aimed at equipping students with advanced knowledge in their field of study but also at fostering cross-curricular competencies (European Commission, 2018; UNESCO, 2012). Among others, the ability to self-regulate one's learning and the (further) development of one's scientific interest are desirable outcomes of higher education (Helmke & Schrader, 2010; Spinath & Seifried, 2018). To achieve these multiple and complex aims of higher education, it is not only university students who are challenged but university teachers who are challenged as well.

On the teachers' side, the challenge is to give as many students as possible the opportunity to learn as much as possible. To do so in an effective way, teachers should offer learning activities that have been found to be effective for learning in empirical studies (i.e., engage in evidence-based teaching; see, e.g., Saville, 2010; Schwartz & Gurung, 2012). On the students' side, the challenge lies in the self-regulated use of learning opportunities in a context that comes with less external control than in secondary school. To do so, students need motivation but also self-regulatory skills and metacognition (Pintrich, 2004). However, empirical studies have found that students' motivation decreases over time (e.g., Benden & Lauermann, 2021; Kosovich et al., 2017), and that students' abilities to correctly monitor and self-regulate their learning are suboptimal (e.g., Bensley et al., 2015; Foster et al., 2017; Peverly et al., 2003). Finally, teaching and learning alike are tied to an institutional context and associated requirements. Such requirements are reflected in, for example, the number of students in courses, the resources teachers have for their teaching, and their capacities to give students individual feedback. The importance of the learning context became even more visible during the COVID-19 pandemic and the online teaching and learning that ensued. This change of context additionally challenged teachers to spontaneously adapt their teaching and challenged learners to self-regulate their learning even more (Lockl et al., 2021; Schwab et al., 2022).

This dissertation builds on research on effective learning activities and multiple desirable outcomes of teaching in higher education. I investigated how students learn and which variables explain several outcomes in self-regulated learning contexts. The focus lies on students' use of specific evidence-based learning activities as a behavioral and proximal predictor of their success in a real-world learning setting.

I begin by introducing a supply-use model for teaching and learning in higher education. This model serves as the theoretical framework for this dissertation and outlines the processes under study. Subsequently, I provide an explanation of the model's key components: the implications of the need for self-regulation in the context of higher education, the various learning outcomes that are desired in self-regulated learning environments, the significance of individual learning prerequisites, the role of motivation as a distinct prerequisite, and, finally, the impact of specific behavior on students' learning outcomes. I describe the specific context and learning activities under study and present open research questions and the aims of this dissertation, followed by the three empirical studies that were conducted to contribute to the understanding of higher education students' learning behavior and motivation in a real-world learning setting. In the empirical studies, evidence-based learning activities were implemented in a lecture class and evaluated across five cohorts of undergraduate students. In the first paper, I and my coauthors examined whether the use of evidence-based learning activities could account for knowledge acquisition beyond learning prerequisites and whether motivation played a role in explaining the use of these activities (Paper 1: Bosch et al., 2021). Then, in light of the challenges posed by the COVID-19 pandemic on student motivation, we compared motivational trends between a regular semester and a pandemic-related online semester and investigated whether evidence-based learning activities could help sustain students' motivation in different contexts (Paper 2: Bosch & Spinath, 2023a). Finally, we examined whether the use of evidence-based activities could account for students' enhanced metacognitive accuracy. Additionally, we explored the possibility of identifying distinct subgroups of students on the basis of their specific patterns of activity use (Paper 3: Bosch & Spinath, 2023b). At the end of this dissertation, I generally discuss the empirical findings, including the strengths and limitations of the three papers. I present directions for future research as well as implications regarding the study of evidence-based learning activities and their practical implementation.

2. Teaching and Learning as Supply and Use

Supply-use models offer an integrative framework for capturing teaching and learning processes as well as different contextual characteristics that are important for learning (Seidel, 2014; Vieluf et al., 2020). Teaching can be understood as the creation of learning opportunities, and successful learning requires the use of these opportunities. Consequently, supply-use models differentiate three levels: the *supply* of learning opportunities, the *use* of learning opportunities, and the learning *outcomes* that result from use of learning opportunities (Brühwiler & Blatchford, 2011; Fend, 1984; Helmke & Weinert, 1997; Seidel, 2014). With supply-use models, the complex interplay of many different factors on different levels in specific learning contexts can be considered (Brühwiler & Blatchford, 2011; Seidel, 2014). However, empirical research that has attempted to simultaneously capture variables from different levels of these models and evaluate their interrelationships is scarce (see also Seidel, 2014).

One basic assumption in supply-use models is that the best learning outcomes can be expected if the supply is of high quality and used to its maximum extent (Fend, 1984). Consequently, teaching quality affects students' achievement indirectly but only if students use the provided learning opportunities (Fend, 1984; Halpern & Hakel, 2003). In a recent review, Christ et al. (2022) evaluated students' learning processes as mediators of teaching quality and students' achievement. The authors found mediating effects, especially when students' behavioral engagement was considered as a mediator.

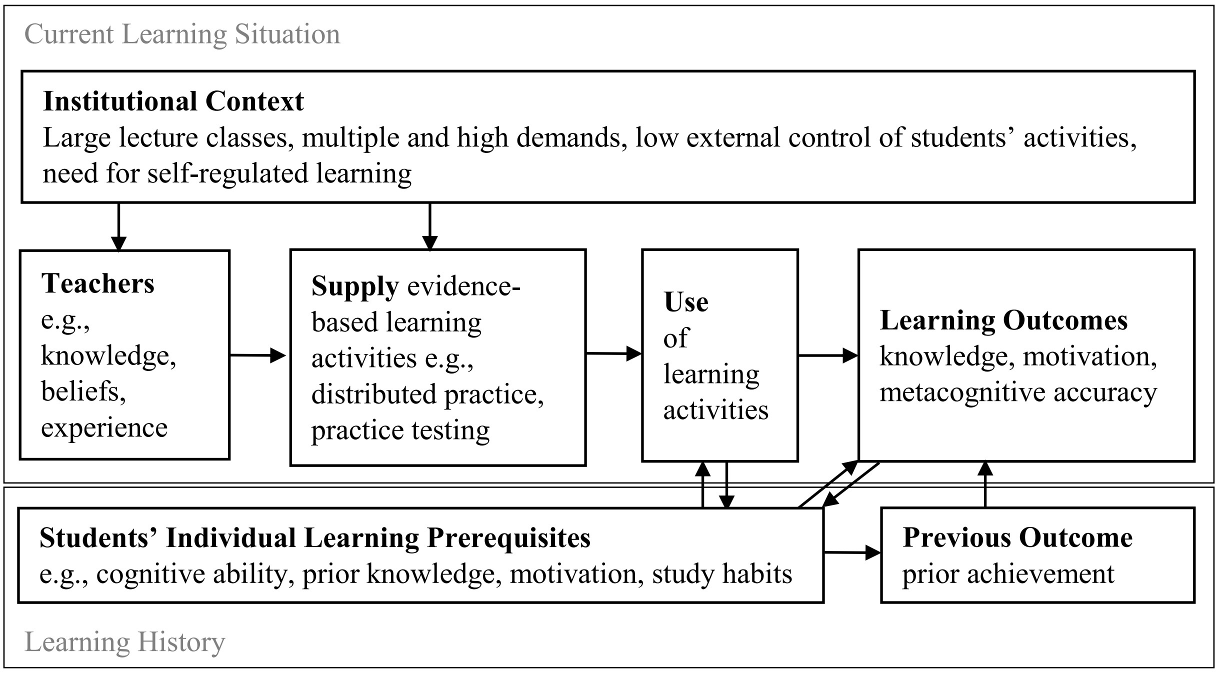
The supply and use of learning opportunities might vary between contexts. Supply-use models consider the context of the learning situation, but so far, they have mostly been applied to the school context. When the context (e.g., university settings) leaves many degrees of freedom to the learners, and the use of activities is optional, the learners' responsibility for their own success increases. Students experience less external control and must decide for themselves which learning opportunities to seize (e.g., Schneider & Preckel, 2017; Seifried et al., 2018; Spinath & Seifried, 2018). As such, learners are considered active participants in their own learning (e.g., Zimmerman, 2002). To determine the predictors of students' active use in contexts with low external control, supply-use models consider individual characteristics, such as students' learning prerequisites (e.g., prior knowledge, cognitive ability, learning strategies, motivation, effort). These prerequisites are thought to explain the use of activities and thereby also indirectly influence learning outcomes via behavior (Brühwiler & Blatchford, 2011; Helmke & Brühwiler, 2018; Helmke & Weinert, 1997; Seidel, 2014).

Building on these models, a supply-use model for university learning and teaching was adopted as the framework for this dissertation (see Figure 1). The basic idea of supply and use in learning situations was maintained. However, in a university setting, further differences compared to school must be considered. Students in higher education have a longer learning history, more knowledge, and more experience with learning from school. Given students' advanced knowledge and experience in education, it could be argued that quality of provided learning opportunities no longer has as great an impact as it did in primary and secondary school. Students' advanced individual learning prerequisites might help them learn regardless of teaching methods and learning opportunities (see Schneider & Preckel, 2017 for a similar reasoning). However, in their review of meta-analyses Schneider and Preckel (2017) identified many variables related to teaching methods with medium to large effects on higher education students' performance. Consequently, although indeed, learning outcomes heavily rely on students' learning prerequisites and learning behavior (Credé & Kuncel, 2008; Köller, 2012;

Richardson et al., 2012), teachers are also challenged to supply effective learning opportunities to achieve the aims of higher education (see, e.g., Boser et al., 2017; Dunn et al., 2013).

Figure 1

A Supply-Use Model of Learning and Teaching at a University (adapted from Bosch et al., 2021)



In the institutional context, many factors can be considered, such as the subject of study, the learning material, and the form of the course (see, e.g., Helmke & Schrader, 2010). Learning institutions define the resources for teaching and sometimes the teaching format. At the undergraduate level at universities, such requirements often result in large lecture classes with many students (Schneider & Preckel, 2017), weekly lecture sessions taught by a lecturer, and an exam at the end of a course (e.g., Rutkiene & Tandzegolskiene, 2015; Watts & Schaur, 2011). Within these limits, university teachers have some degrees of freedom to design their courses (e.g., how to present the content or which additional activities and feedback to offer). In large courses, teachers should offer learning opportunities that can be used by many students and will enable students with heterogeneous learning prerequisites to learn as much as possible. To achieve this goal, teachers should engage in evidence-based teaching and should supply their students with learning opportunities that have been shown to promote student learning in controlled settings (Boser et al., 2017; Dunn et al., 2013). Considering the supply of learning opportunities as a central aspect of teaching focuses on changeable aspects of teaching and consequently enables university teaching to be improved for different teachers (Spinath & Seifried, 2018).

2.1 Self-Regulated Learning in Higher Education

The self-regulation of learning is required for students to be successful when studying at the university level. From the self-regulated learning perspective, "learning is viewed as an activity that students do for themselves in a proactive way rather than as a covert event that happens to them in reaction to teaching" (Zimmerman, 2002, p. 65). Learners are conceptualized as metacognitively, motivationally, and behaviorally active participants in their own learning (Zimmerman, 1986). Connecting this thought to supply-use models implies that students need to regulate their participation in the activities that are offered to them.

The learning process itself is described as a cyclical process in several models of self-regulated learning, including feedback loops (e.g., Schmitz & Wiese, 2006; Winne, 1996; Zimmerman, 2002; 2008). The models describe three cyclical phases: In the *preaction* phase, learners analyze the task (in relation to their needs, abilities, and motivation), plan activities, and set goals. In the *action* phase, actions are taken to learn. Learners need to control their learning strategies and monitor their own progress. In the *self-reflection* phase, learners evaluate learning outcomes and react to the results. The outcome of the self-reflection phase may shape how the next cycle of learning unfolds. This cyclical process as well as reciprocal effects are also indicated in Figure 1. For example, when considering the role of motivation in self-regulated learning in general (Schwinger & Stiensmeier-Pelster, 2012) and as a potential predictor of ongoing activity use more specifically (e.g., Putwain et al., 2019), the maintenance of motivation is an important learning prerequisite to use learning activities, but it is also an outcome of activity-use. Further, a learning outcome (e.g., acquired knowledge) from the first application of a learning activity might be integrated as a new learning prerequisite (e.g., prior knowledge) for a different future activity. And if students have good experiences with a learning activity, they might add it to their repertoire and improve their study habits.

In an ideal scenario, students are aware of their own strengths and limitations, set their goals realistically, choose appropriate strategies, monitor their behavior and progress, evaluate their progress accurately, and then adapt their goals and strategies accordingly (see Zimmerman, 2002). This scenario highlights the need for students to accurately monitor their own learning process and outcomes. Only when students monitor themselves accurately can they adapt their goals and strategies in a goal-directed way. In addition, self-regulated learners might experience self-satisfaction and motivation when they evaluate their increasing competence and effectiveness throughout the learning process (Zimmerman, 2002). However, not all students attending a university are ideal self-regulated learners (e.g., Bensley et al., 2015; Peverly et al., 2003), and, for example, university students' learning prerequisites are becoming more diverse as more students are earning a high school diploma (Bernhard, 2021; Eckert et al., 2015; Organisation for Economic Co-operation and Development [OECD], 2018; 2019). As a consequence, self-regulation processes differ between individuals, and many students struggle with self-regulation. This phenomenon can be seen in intention-behavior gaps regarding students' own learning behavior (Dobronyi et al., 2019), negative changes in motivation over time (e.g., Benden & Lauermann, 2021; Kosovich et al., 2017), heightened procrastination among students (Steel & Ferrari, 2013), as well as an inaccurate monitoring of students' own knowledge (e.g., Bensley et al., 2015; Dunning et al., 2003; Foster et al., 2017).

2.2 Desirable Learning Outcomes

Although prior learning outcomes (e.g., knowledge and self-assessment) are incorporated into the next learning cycle as learning prerequisites (e.g., Zimmerman, 2008), it is helpful to differentiate learning prerequisites from learning outcomes within one learning situation so that both can be investigated empirically. A central goal of higher education is to impart knowledge, but it is also important for university students to develop cognitive, volitional, emotional, and social competencies (European Commission, 2018; see also Helmke & Schrader, 2010). Higher education should therefore tackle cognitive, motivational, affective, and behavioral aspects (Spinath & Seifried, 2018). A large body of research has focused on knowledge and performance as central learning outcomes in higher education courses (see Schneider & Preckel, 2017, for an extensive review of meta-analyses explaining performance), and it is intuitively reasonable to assume that students should acquire knowledge. In this dissertation, I investigated three learning outcomes. First, I investigated the *knowledge* that students acquire in a lecture class. Second, I also investigated *maintained motivation* as a desirable learning outcome. The maintenance of motivation is important for students to initiate learning processes repeatedly (e.g., Pintrich, 2004; Putwain et al., 2019), and the need to maintain motivation is even more pronounced in the absence of external control and the need to learn continuously over the course of one academic term (see Section 2). In addition, when students leave higher education courses with interest in their field of study and the motivation to learn more, such an outcome matches higher education's goal of fostering further interest (Helmke & Schrader, 2010). Finally, in the literature on self-regulated learning, metacognition is considered an important learning outcome (see Section 2.1). Students need to become aware of their own knowledge in order to initiate the next learning cycle in an adaptive way (see, e.g., Schmitz & Wiese, 2006). Therefore, I explored *metacognitive accuracy* as a third desirable learning outcome in this dissertation (see also Thiede et al., 2003). Enhanced metacognition and maintained motivation also tackle the cognitive and motivational aspects of learning mentioned by Spinath and Seifried (2018).

2.3 University Students' Learning Prerequisites

Individual learning prerequisites are understood as all the knowledge, skills, experiences, and characteristics students bring with them to the current learning situation. Across the 38 meta-analyses reviewed by Schneider and Preckel (2017), 63 student variables were considered to explain achievement, of which 32 could be considered individual learning prerequisites. The authors listed motivation, intelligence, and prior achievement among the 10 most important predictors of achievement (d = 0.79–1.81). In secondary school, approximately 50 % of the variance in academic achievement is explained by individual learning prerequisites (Helmke & Schrader, 2010; Köller, 2012), but how much variance in university achievement is explained by individual learning prerequisites is less clear. In a meta-analysis, Richardson et al. (2012) found that prior achievement, along with conscientiousness, effort regulation, test anxiety, and academic self-efficacy, explained 28 % of the variance in college grade point average.

One of the most important single predictors of school achievement is cognitive ability (Rost, 2009). In higher education, students have successfully mastered secondary education and are therefore believed to have–on average–better learning prerequisites (e.g., intelligence) than the general population (Credé & Kuncel, 2008; Sackett et al., 2009). Furnham et al. (2002) contended that in higher education, variance in intelligence tends to decrease, particularly in selective fields of study where admission criteria typically admit only high-achieving students with high cognitive abilities. Although intelligence is still an important predictor of achievement in higher education (Busato et al., 2000; Lounsbury et al., 2003; Rohde & Thompson, 2007), it is essential to look for further explanatory student variables in higher education (see also Richardson et al., 2012). Several empirical studies have additionally identified prior knowledge as important individual predictors of learning success in higher education (Dochy et al., 1999; Thompson & Zamboanga, 2004). Students with high prior knowledge can build on their existing knowledge and are therefore likely to perform better on exams. Furthermore, prior knowledge in part results from prior effort, which was found to be quite stable over time (Gottfried et al., 2001). Schneider and Preckel (2017) reported a large effect size for the relationship between prior knowledge and achievement in higher education ($d = 0.79$).

Prior achievement is the result of many different prior learning situations, which also include intelligence, prior knowledge, and motivation. This cumulative effect might explain why prior achievement is a powerful predictor of university achievement. For example, high school grade point average (GPA) is a good predictor of success at university ($r = 0.36$, Kobrin et al., 2008; $\rho = 0.40$, Richardson et al., 2012; ß = 0.29, Robbins et al., 2004; $d = 0.90$, Schneider & Preckel, 2017). Consequently, prior knowledge and prior achievement should be included as important control variables when investigating other important determinants of success. When controlling for prior knowledge and achievement, variables such as motivation, study habits, and skills still remain important individual difference variables that explain achievement (Credé & Kuncel, 2008; Richardson et al., 2012).

3. Learning Motivation

Motivation plays a crucial role in students' achievement, as it is thought to initiate and direct learning behavior (Credé & Kucel, 2008; Wigfield et al., 2015). The importance of motivation beyond cognitive prerequisites and prior achievement has been shown in numerous empirical studies ranging from the secondary school context (e.g., Dickhäuser et al., 2011; 2016; Kriegbaum et al., 2018) to higher education (Credé & Kuncel, 2008; Eckert et al., 2015; Richardson et al., 2012; Robbins et al., 2004; Schneider & Preckel, 2017).

In this dissertation, motivation is conceptualized in accordance with Expectancy Value Theory (EVT; Eccles et al., 1983; Wigfield & Eccles, 2000). According to EVT, expectancies (how well students think they will perform) and values (how important, enjoyable, and useful a subject or task is for them) explain students' achievement-related choices and behavior, which in turn explain performance. These assumptions have received a great deal of empirical support: When explaining performance in higher education, the expectancy-related component of motivation is especially important to consider (Richardson et al., 2012; $d = 1.81$ for performance self-efficacy, Schneider & Preckel, 2017), whereas values explain achievement-related choices and engagement in a subject (Musu-Gillette et al., 2015). Applying this differentiation to the supply-use model (see Section 2), it might be plausible that expectancies are more closely related to acquired knowledge, whereas values might explain use of learning activities.

If students are motivated to learn, this is associated with better performance (Lazowski & Hulleman, 2016; Robbins et al., 2004). However, higher education students' motivation has been found to decrease over time (e.g., Benden & Lauermann, 2021; Kosovich et al., 2017; Musu-Gillette et al., 2015; Robinson et al., 2019). In the worst-case scenario, students fail to meet their academic goals and drop courses (Benden & Lauermann, 2021) or even quit their whole study program (Heublein & Wolter, 2011) due to insufficient motivation. Considering the need for self-regulated learning in higher education, a loss of motivation appears especially problematic. When students make independent decisions about their learning activities amidst numerous demands and choices, as it is often the case (see Pintrich, 2004), motivation is required continuously but can also be subject to threats. Situated EVT (Eccles & Wigfield, 2020) considers expectancies and values to be situationally sensitive. The sensitivity of motivation to specific situations may have resulted in additional challenges in the context of online learning. Due to the lockdowns prompted by the COVID-19 pandemic in March 2020, both students and teachers had to adapt to new teaching and learning formats. Some empirical studies found evidence that these new ways of teaching and learning had a negative impact on students' motivation (Garris & Fleck, 2022; Müller et al., 2021; Usher et al., 2021) and increased the demands on self-regulated learning (Lockl et al., 2021).

Although mostly considered a learning prerequisite, motivation is also a desirable learning outcome at universities (Helmke & Schrader, 2010). To understand the development of motivation, EVT (Eccles & Wigfield, 2020) has assumed cyclical processes: Previous achievement-related experiences are believed to shape motivational values. Also, in self-regulated learning models, each action taken for learning is followed by a reflection on its outcomes. This reflection then contributes to building up the motivation to engage in subsequent learning-related actions (Schmitz & Wiese, 2006). Additionally, Hidi and Renninger (2006) contended that eliciting situational interest can encourage repeated engagement, ultimately strengthening students' overall interest in a particular domain. Consequently, students' learning behavior, particularly their engagement in evidence-based learning activities, is considered one potential means for generating situational interest and influencing achievement-related experiences, which, in turn, could shape future motivational values. However, empirical research investigating behavior as an antecedent of motivation is scarce.

Taken together, motivation is an important learning prerequisite, but empirical studies have indicated a negative trend in students' motivational development over time. At the same time, motivation can be considered a desirable learning outcome. Research is needed to examine the learning activities that are offered and used in the context of self-regulated learning and their potential relationships with motivation as well as additional learning outcomes, including knowledge and metacognitive accuracy.

4. Investigating Learning Behavior

"What professors do in their classes matters far less than what they ask *students* to do" (Halpern & Hakel, 2003, p. 41). Although provocatively phrased, this quote emphasizes once more the important message that students play an active role in the learning process and that the teachers' role is primarily to provide learning opportunities and get students to seize them. Knowledge and other desirable learning outcomes are not passively accumulated by students in reaction to teaching but need to be actively built by the students, thus requiring considerable behavioral and cognitive activity (Helmke & Schrader, 2010). In models of self-regulated learning, this process is depicted in an explicit *action phase* (see Section 2.1); and in the supply-use model, the active utilization of learning activities is considered a proximal determinant of learning outcomes (see Section 2 and Figure 1).

Accordingly, in a meta-analysis, Richardson et al. (2012) found that effort regulation was the only learning strategy that explained performance when prior achievement was controlled for ($\Delta R^2 = .03$). Credé and Kuncel (2008) found that study habits (study routines; e.g., the frequency of study sessions, self-testing, review of material) explained variance in achievement in college after they controlled for prior achievement ($\Delta R^2 = .03$). In a recent review, Christ et al. (2022) investigated the mediating role of learning processes between teaching quality and students' achievement. They found stronger mediating effects when engagement was investigated as behavioral or metacognitive engagement rather than as motivational beliefs or goal setting strategies. Further, in a systematic review and meta-analysis on students' engagement, Wong et al. (2023) found that studies that considered a behavioral component of engagement found larger effects on academic achievement ($r = .39$) than studies that considered cognitive ($r = .31$) or affective ($r = .26$) engagement measures. These two systematic reviews further emphasized the usefulness of studying effort or engagement when explaining learning success. Further, it seems more useful to assess actual behavior rather than using self-report scales that assess the use of abstract strategies by asking participants to indicate which strategies they typically apply across learning situations (e.g., Pintrich et al., 1993; Richardson et al., 2012). Students' learning behavior seems to be a promising predictor of achievement. But still, behavioral components (Wong et al., 2023), study habits (Credé & Kucel, 2008), engagement (Christ et al., 2022), and effort (Richardson et al., 2012) are all rather vague terms that need to be operationalized specifically in empirical studies.

4.1 Evidence-Based Learning Activities

To conduct more accurate investigations of what exactly students do when they invest time, effort, or engagement, students' learning behavior should be assessed by targeting *specific* activities in a *specific* learning situation (see also Artelt, 2000). When students need to organize their learning on their own, they do not always use the most effective learning activities (see, e.g., Blasiman et al., 2017; Karpicke et al., 2009; Kornell & Bjork, 2007). However, teachers can offer evidence-based learning opportunities that have been shown to promote learning in empirical studies (Saville, 2010; Schwartz & Gurung, 2012). Researchers are calling for teachers to engage in evidence-based teaching at universities because such teaching is thought to enhance not only students' knowledge acquisition but also further desirable outcomes in higher education, such as metacognition, long-term retention, and transfer (e.g., Dunn et al., 2013; Dutke et al., 2017; Halpern & Hakel, 2003; Slavin, 2008). When teachers offer evidence-based activities, students' learning can be optimized (Boser et al., 2017; Dunn et al., 2013; Pashler et al., 2007).

Learning activities that help students distribute their learning (Dunlosky et al., 2013) and test themselves repeatedly (Bae et al., 2019; Butler & Roediger, 2007; Cogliano, et al., 2019; Dunlosky et al., 2013; Jonsson et al., 2021; Roediger & Butler, 2011; Rowland, 2014) have been shown to contribute to performance. Further, if students receive continuous feedback on their current state of knowledge, they can adapt their learning accordingly (Butler et al., 2007; Downs, 2015; Janson et al., 2023). In this line of thinking, for example, Leeming (2002) found that students learned more in a lecture class when short tests were administered at the beginning of each class. Practice testing has been found to improve not only students' knowledge but also their metacognitive accuracy because practice tests enhance memory and ease future retrieval (e.g., Barenberg & Dutke, 2019; Carpenter, 2012; Dunn et al., 2013; Rowland, 2014; Thomas & McDaniel, 2013). The beneficial effects of testing on learning have been investigated empirically many times (for a meta-analysis of over 200 classroom studies, see Yang et al., 2021; effect of testing on performance Hedges' $g = 0.49$). Further research has shown that writing to learn is helpful for students' performance because it stimulates additional cognitive processes and improves comprehension (for a meta-analytic review, see Schindler & Richter, 2023; Hedges' $g = 0.41$). Regular attendance of classes has also been shown to relate to performance (for a meta-analysis, see Credé et al., 2010; $\rho = 0.44$). Especially when lecture attendance is combined with independent review of the material, students can distribute their practice and improve their performance (Cepeda et al., 2006; see also Credé et al., 2010). However, research on evidence-based learning activities is oftentimes focused on testing and its effects on cognitive outcomes. Empirical research investigating multiple optional learning activities is scarce (see Naujoks et al., 2022). On the basis of this empirical evidence, specific activities were implemented in a lecture class and investigated in this dissertation with regard to multiple learning outcomes.

4.2 The Context of this Dissertation: Investigating a Lecture Class

Empirical research on effective teaching is particularly needed in the higher education sector (Spinath & Seifried, 2018). Halpern and Hakel (2003) described that, despite a desire among university faculty to provide good teaching, and though they may have a great deal of subject knowledge, they often know little about learning or memory. As a result, many teachers tend to fall back on intuitive knowledge about good teaching without systematically examining it. Applying research methods to one's own teaching is one way to improve teaching (Scholarship of Teaching and Learning; e.g., Hutchings et al., 2011; see also research-based teaching, Spinath et al., 2014). Engaging in research-based teaching involves an ongoing process where didactic actions are chosen on the basis of theory and empirical data, implemented, and tested for effectiveness using appropriate research designs to enhance teaching and its empirical foundation (Spinath & Seifried, 2012). Accordingly, optional evidence-based learning activities were chosen on the basis of empirical data regarding effective learning activities (see Section 4.1), and implemented in a lecture class. In the next step, they were tested for effectiveness in this dissertation. The aims of the lecture class were to provide undergraduate psychology students and preservice teacher students with basic knowledge about important topics in educational psychology (e.g., how to improve learning). Furthermore, as the program was given rather early in the curriculum, and evidence-based acting and thinking should be fostered, the lecture class was also dedicated to providing students with basic statistical knowledge to enable them to interpret empirical evidence regarding topics in educational psychology (e.g., to understand the concept of correlations and the importance of conducting experiments in order to draw causal conclusions).

The studied lecture class in this dissertation was a typical learning situation at the undergraduate level (see Section 2; Rutkiene & Tandzegolskiene, 2015; Schneider & Preckel, 2017; Watts & Schaur, 2011): a large introductory lecture class on educational psychology with weekly lecture sessions and an exam at the end of the semester. Besides weekly sessions, the following learning activities were offered across all cohorts from spring 2018 to spring 2022.

On the basis of empirical evidence for practice testing (see Section 4.1), brief knowledge tests with confidence-weighted true/false items (Dutke & Barenberg, 2015) were included five times over the course of one semester (i.e., every three weeks) on which students received corrective feedback (their individual scores as well as the correct solutions to the items). Further, the opportunities to submit essays addressing thought-provoking questions and to apply the contents of the lecture to practical problems were offered eight times over the course of the semester (see also Schindler & Richter, 2023; Seifried et al., 2018). Due to the importance of feedback (see, e.g., Downs, 2015), student tutors were trained to give individual criterion-based feedback on these submissions. By offering these activities with feedback multiple times, students were supported in distributing their practice, continuously increasing their own knowledge, and achieving an accurate assessment of their own knowledge (formative self-assessment; see also Seifried & Spinath, 2021). In addition to these activities, frequently attending classes and reviewing materials repeatedly was also suggested to improve knowledge, as it has been shown to be related to performance (see Credé et al., 2010; Section 4.1). Frequent use of these activities can also help students space out their learning and was therefore considered effective (see also Dunlosky et al., 2013). Use of learning activities was voluntary for all students and activity use was assessed anonymously.

After implementing evidence-based learning activities (for a detailed lecture class description, see also Seifried & Spinath, 2021), this dissertation aimed to empirically evaluate the optional use of these activities and consider both antecedents and outcomes.

5. Open Research Questions Addressed in This Dissertation

Using the supply-use model (Figure 1, p.10), I would like to summarize the most important assumptions and research questions derived for this dissertation. As outlined above, university students have good learning prerequisites on average, but they face new challenges. In the university context, they are required to learn in a self-directed manner to a greater degree than they have previously experienced. Therefore, in order to succeed, they must decide for themselves how to learn, maintain their motivation over time, engage in learning behavior, and evaluate their learning accurately. Teachers can support this process by providing effective learning opportunities even for large lecture classes.

Several evidence-based learning activities have been shown to be effective in enhancing university students' learning (see, e.g., Dunlosky et al., 2013; Dunn et al., 2013). However, it is less clear which learning activities students use when they are continuously offered as optional activities in a lecture class over the course of one semester and how they relate to several outcomes. Empirical research so far has often focused on the effectiveness of single learning activities (e.g., testing; see Yang et al., 2021). The dissertation builds on studies that have investigated the effectiveness of learning activities (e.g., Bae et al., 2019; Credé et al., 2010; Dunlosky et al., 2013; Händel et al., 2020; Naujoks et al., 2022; Schindler & Richter,

2023) and extends them by investigating the *supply* of more optional learning activities all together in one lecture class and evaluating how much students *use* them. So far, there is evidence that students use a mixture of ineffective and effective strategies to prepare for exams (Blasiman et al., 2017). Consequently, the first aim of this dissertation was to evaluate students' activity use. Which activities do they use (the most) and to what extent? As activity use was expected to differ interindividually, another open question was whether subgroups of students with specific patterns of learning activity use exist.

After offering evidence-based activities, of course, their effectiveness should be tested (see Spinath & Seifried, 2012). According to the model, the use of learning opportunities is considered a proximal predictor of learning outcomes. In a university setting, several *outcomes* can be considered desirable. First, it is a given that knowledge should be acquired in university courses. However, maintained motivation and an accurate self-assessment of their own knowledge (metacognition) are important for students to succeed while attending university courses (e.g., Pintrich, 2004; Thiede et al., 2003). Empirical studies have investigated the effect of testing on knowledge and metacognition (e.g., Barenberg & Dutke, 2019; Naujoks et al., 2022). However, it is less clear how other learning activities contribute to these outcomes when they are all offered together and how active participation in learning activities might help students sustain their motivation. Especially under threatened conditions, such as emergency distance learning due to the COVID-19 pandemic, activities that can easily be implemented even in online teaching and that may help maintain motivation when it is under threat seem promising. Consequently, a second aim of the dissertation was to explain these diverse learning outcomes. Does the use of evidence-based activities explain knowledge acquisition, motivation, and metacognitive accuracy beyond individual learning prerequisites and prior achievement?

Last but not least, predictors of activity use were investigated. In supply-use models, *individual learning prerequisites* are believed to explain the use of learning opportunities, i.e., offered learning activities, as well as learning outcomes (see Figure 1). It was therefore investigated, how interindividual differences in activity use can be explained.

6. Summaries of Empirical Studies Included in This Dissertation

In the following, I present the findings of three empirical studies that were conducted to address the open research questions outlined above. The published versions of the three papers are included at the end of this dissertation.

In all three studies, students' use of learning activities over the course of one semester was examined. In each study, the benefits of evidence-based learning activities in an introductory lecture class on educational psychology were investigated, each time focusing on one desirable outcome of higher education learning: knowledge, motivation, and an accurate self-assessment. The lecture class included weekly lecture sessions. Besides these sessions, optional learning activities with regard to evidence-based principles were offered (see Section 4.2).

6.1 Learning Activities Explain Knowledge Acquisition (Paper 1; Bosch et al., 2021)

In Paper 1, we examined students' learning behavior by looking at their learning intentions and use of specific evidence-based learning activities in the lecture class. In addition, we conducted the first investigation of whether learning activity use contributed to students' knowledge acquisition beyond their learning prerequisites and prior achievement. We assumed that, beyond more distal antecedents of students' achievement (e.g., learning prerequisites), their specific behavior in a specific learning situation would contribute to explaining their knowledge acquisition (see also the supply-use model; Figure 1). To investigate these assumptions, we assessed $N = 112$ students' learning prerequisites (prior knowledge and motivation) and prior achievement (high school grade point average) as well as their intentions to use the learning activities at the beginning of a lecture class in spring 2018. At the end of the course, we assessed students' use of learning activities as well as their knowledge in a mock exam.

We found that students actually used significantly fewer activities than they had intended to at the beginning (d = 1.61). However, when students used learning activities, this use explained their knowledge acquisition beyond their learning prerequisites and prior achievement in hierarchical regression analyses (ß = 0.20, $p < .05$; $R^2 = .17$ in the full model, $p < .01$). In the next semester, we were able to replicate these findings in a second sample ($N = 171$). In bivariate associations, several learning activities were associated with knowledge at the end of the semester (e.g., lecture attendance, submission of essays, self-directed reviewing of the lecture slides and notes). We further investigated the predictors of activity use. Motivational values at the beginning of the semester explained the use of learning activities during the semester (ß = 0.29, $p < .01$), whereas expectancies did not. On the one hand, the results showed that university students failed to use activities to the extent that they had intended to. On the other hand, knowledge acquisition was explained by activity use beyond learning prerequisites and prior achievement, strengthening the empirical evidence for the effectiveness of some activities and emphasizing that students had opportunities to succeed by being active participants in their learning process.

6.2 Learning Activities Explain the Development of Motivation (Paper 2; Bosch & Spinath, 2023a)

In Paper 1, we found that students who valued the subject more, engaged in more learning activities. However, in models of EVT, previous achievement-related experiences also affect subjective task value (Eccles & Wigfield, 2020). In addition, a specific requirement for students in higher education is to maintain their motivation for ongoing self-directed learning, but empirical studies have found negative trends in students' motivational values (e.g., Benden & Lauerman, 2021, Kosovich et al., 2017). Therefore, in Paper 2, we focused on the development of intrinsic value and whether learning activity use was related to it.

Furthermore, in March 2020, universities had to close, and students had to learn from home on relatively short notice, a situation that increased the demands on students and teachers (Lockl et al., 2021; Schwab et al., 2022). Initially, it was feared not only that less learning would take place but also that students' motivation would develop negatively (Usher et al., 2021). To test these ideas, students' initial motivation and motivational development were compared between regular and online semesters. Less motivation at the beginning of the semester and a more pronounced negative development of motivation were hypothesized for students attending an online course during the pandemic. However, regardless of the format (online or in-person), students' motivation was expected to benefit from the use of evidence-based learning activities. Over the course of the semester, we assessed students' intrinsic value for the subject five times in two cohorts of students (summer term 2019, $n = 121$, in a regular lecture class; then again, during the summer term of 2020, $n = 119$, after the shift to distance learning). At the end of the semester, students also indicated how many learning activities they had used.

After confirming strong measurement invariance for the assessment of intrinsic value, we computed a multigroup second-order linear change model. We found that students' motivation declined in both cohorts, but we did not find significant differences in students' motivation between cohorts (Change Factor $M = -0.29$, 95 % CI [–0.43, –0.16] in 2019; $M = -0.32$, 95 % CI [–0.44, –0.21] in 2020). Further, in both cohorts, students' use of learning activities was positively associated with change in intrinsic value, even after controlling for initial intrinsic value, indicating that the decline in students' motivation was not as steep for students who engaged in many learning activities ($r_{2019} = .49$, $p < .01$; $r_{2020} = .45$, $p < .01$).

The results emphasized another advantage of offering evidence-based learning activities. Although we could not draw causal conclusions due to the study design, the use of learning activities was associated with a less steep decline in motivation over the course of one semester—regardless of the format of the lecture class. Learning activities that are easy to implement online, in particular, could offer an opportunity to facilitate learning for students, even in distance learning.

6.3 Learning Activities Explain Metacognition (Paper 3; Bosch & Spinath, 2023b)

In Paper 3, we investigated students' activity use in more detail. When attending a lecture class, it is plausible that students use several learning activities and combine them individually. Therefore, we investigated individual patterns in learning activity use in two large educational psychology lecture classes. In Paper 3, we further aimed to replicate the Paper 1 finding that the use of learning activities explained knowledge acquisition beyond individual learning prerequisites and prior achievement. In addition to knowledge acquisition, we considered indicators of students' metacognitive accuracy (awareness of their own content specific strengths and weaknesses, i.e., correct confidence, as well as an accurate assessment of their own knowledge level, i.e., accuracy) and how they change over the course of one semester. In empirical studies, students often tend to overestimate their level of knowledge and are overconfident (Bensley et al., 2015; Dunning et al., 2003; Foster et al., 2017; Hartwig & Dunlosky, 2014; Händel & Dresel, 2018). In the winter term 2021, we assessed $N = 285$ students' individual learning prerequisites at the beginning of the course and learning outcomes at the end of the course, again. In this study, the use of activities was continuously assessed over the course of the semester. Throughout the semester, students reported their use of each learning activity four times. Using latent profile analysis, we identified five different profiles of activity use. Whereas most students showed high, average, or low use of all activities, some students focused on reviewing notes and participating in knowledge tests (5 % of the students), whereas others focused on attending the lecture and submitting essays (6 %).

The results showed that students' knowledge and correct confidence increased over the course of one semester ($d = 1.81$ and $d = 0.84$, respectively) and that students benefitted from the frequent use of many different activities. In multivariate regression analyses, the mean use of learning activities explained variance in knowledge as well as correct confidence and accuracy beyond students' learning prerequisites and prior achievement ($\Delta R^2 = .16$ for knowledge; $\Delta R^2 = .05$ for correct confidence; $\Delta R^2 = .04$ for accuracy). Comparing learning outcomes between groups from the latent profile analysis also suggested that students benefitted the most when they had used many different activities frequently.

Whereas the results regarding the latent profiles must be considered exploratory, the associations between overall activity use and the outcomes reinforced the effectiveness of evidence-based learning activities in a lecture class. Students who engaged in more activities continuously acquired more knowledge as well as greater metacognitive accuracy.

7. General Discussion

Due to the manifold aims of higher education (Helmke & Schrader, 2010; European Commission, 2018; Spinath & Seifried, 2018) as well as the call to design and improve university teaching on the basis of empirical evidence (Dunn et al., 2013; Hutchings et al., 2011; Spinath & Seifried, 2012), the overarching aim of this dissertation was to evaluate the use and usefulness of evidence-based learning activities in a lecture class. Emphasizing students' active role in the learning process (Credé & Kuncel, 2008; Halpern & Hakel, 2003), the use of optional activities was assessed in a real-world learning setting. Considering empirical findings that have emphasized the need for students to be engaged or to show effort in order to succeed (Christ et al., 2022; Richardson et al., 2012; Wong et al., 2023) as well as evidence of the effectiveness of specific learning activities (e.g., practice tests and distributed learning, see e.g., Bae et al., 2019; Cogliano, et al., 2019; Dunlosky et al., 2013; Jonsson et al., 2021; Roediger & Butler, 2011; Rowland, 2014; Yang et al., 2021), the presented studies combined these ideas and investigated the use as well as antecedents and outcomes of multiple evidence-based learning activities in the field. In the following, the results from the three papers are discussed together while also addressing strengths and limitations, ideas for future research, and practical implications.

7.1 Supply and Use of Evidence-Based Learning Activities

In all Papers, evidence-based learning activities were implemented in a real-world learning setting and their use was assessed to capture students' active engagement in learning. Results showed high intercorrelations across use of learning activities (Paper 3). Therefore, mean learning activity use and associations with outcomes could be investigated. In empirical studies, students' learning is often assessed using retrospective self-assessments of abstract strategies for studies in general (e.g., Pintrich et al., 1993). How this relates to performance on a specific task is questionable (Artelt, 2000). In the literature, several methods have been investigated to assess learners' study time and strategies more precisely (Liborius et al., 2019; Molenaar et al., 2021). By assessing use of specific activities, the assessment of students' learning in the field was further improved (see also Artelt, 2000). In addition to taking mean activity use into consideration, all three studies also inspected single learning activity use and associated outcomes in exploratory investigations.

From Paper 1 to Paper 3, we aimed to improve the assessment of activity use by applying more differentiated scales to enable students to correctly indicate their activity use. In Paper 2 and Paper 3, we asked students to give the absolute number of activities they engaged in rather than asking them to indicate their use of activities on a scale from 1 (*not at all*) to 5 (*very much;* see Bosch et al., 2021). In Paper 3, students were also asked every 3 weeks about their activity use rather than only once at the end of the semester. Asking students repeatedly about the frequency of specific behavior in a specific learning situation should result in a more accurate assessment than asking about usual strategy use (Artelt, 2000). Whereas in Paper 1, we aimed to assess students' learning behavior as widely as possible by including eight activities (e.g., being part of a study group or reading additional literature), in Papers 2 and 3, we focused more on the activities that were explicitly offered by the teacher and that were related to achievement in Paper 1 (i.e., lecture class attendance, review of lecture slides, participation in online knowledge tests, and the submission of essays). With this refined focus, we explored learning opportunities that university teachers can supply and modify, making them a suitable approach for enhancing higher education teaching (see also Spinath & Seifried, 2018).

Because the assessment of activitiy use was specific, it allowed for the identification of differences in the use of activities. In all three empirical studies, students used the offered activities to different extents. Across all three studies, students primarily engaged in the activities of attending lectures and taking knowledge tests. For example, approximately 60 % of classes were attended as well as 60 % of knowledge tests were taken, whereas only about 20 % of possible essays were submitted in Paper 3. It further became evident that students' actual use of activities differed between students. Due to the diversity of the offered activities, we assumed that certain activities might be more appealing to some students than others. This led us to explore whether students engaged in specific, individually fitting patterns of activity use (see, e.g., Janson et al., 2023 for thoughts on aptitude-treatment interactions). Therefore, we applied a person-centered approach in Paper 3. With this approach, we were able to differentiate five student subgroups that differed in their use of activities. The results showed that students differed primarily in their overall use of all activities: 89 % of students were assigned to subgroups that reported high, medium, or low levels of participation in all activities. However, 11 % of students applied a specific pattern of activity use, focusing on single activities. For example, 5 % of students focused on individual review of lecture slides and participation in knowledge tests, while they hardly attended lecture classes and submitted below-average numbers of essays. For self-regulation of learning, it might even be adaptive to use the learning activities that were individually most useful (e.g., it is also considered most useful to apply an individually suitable motivation regulation strategy, see Schwinger et al., 2012). However, the sample included in Paper 3 was rather small for latent profile analysis

($N = 285$) so specific activity use patterns across the small groups are interpreted with caution and should be replicated in future studies.

7.2 Explaining Desirable Learning Outcomes

Active use of learning activities is considered a proximal determinant of learning outcomes (see Section 2 and Figure 1). Therefore, the extent to which several learning outcomes could be explained by activities beyond individual learning prerequisites and prior achievement was investigated. As has frequently been shown in the past (Kobrin et al., 2008; Richardson et al., 2012), prior achievement also explained knowledge at the end of a semester in our studies. On the one hand, this finding emphasizes the explanatory power of prior achievement, but it also emphasizes the importance of all variables that explain the acquisition of knowledge beyond prior achievement. In our studies, the use of learning activities explained the acquisition of knowledge beyond individual learning prerequisites and prior achievement in Paper 1 and Paper 3. In regression analyses, especially in Paper 3, when replicating the effect from Paper 1, the effect was rather large ($\Delta R^2 = .16$). Prior knowledge and prior achievement together with use of learning activities explained 44% of variance in knowledge at the end of the semester. The comparison of student subgroups in Paper 3 also showed the greatest increase in knowledge in the group of students who participated in all learning activities to a high degree. The results support the idea that students' active engagement matters and that they can actively make a difference in their own success.

These results complement prior empirical studies that have found learning strategies, beyond effort regulation, to be unrelated to students' achievement after controlling for prior achievement (e.g., Richardson et al., 2012). Other studies have found that students' engagement is predictive of learning success beyond prior achievement, albeit with small effects (e.g., $\Delta R^2 = .03$ Credé & Kuncel, 2008). Our results support the theoretically acknowledged importance of learning processes as a proximal predictor of learning outcomes in supply-use models (Seidel, 2014) and the particular importance of the behavioral component of learning (Christ et al., 2022; Wong et al., 2023).

Furthermore, investigating the effectiveness of specific learning activities has great practical importance for university teaching and learning. By providing detailed descriptions of these activities, it becomes possible for other university teachers to adopt them, fostering evidence-based teaching practices (Saville, 2010; Schwartz & Gurung, 2012). This is particularly important given the lack of obligatory pedagogical training for university teachers (see also Halpern & Hakel, 2003). For learners, the supply of evidence-based learning activities provides good guidance for effective learning in a self-regulated learning context.

The second outcome of interest was students' maintenance of motivation (Paper 2). Again, the more students used the learning activities that were offered, the less their motivation declined over the course of one semester, controlling for initial motivation. Therefore, evidence-based learning activities may be a suitable tool for helping students face their motivational challenges. The associations between learning activities and motivational change were moderate to large in size. Of course, as investigated in Paper 1, motivation is also important for engaging in learning activities. However, the continued engagement may also help students maintain their motivation over time because learning behavior is not only a consequence but also a prerequisite, as many models of motivation include past behavior as an antecedent of motivation (e.g., Eccles & Wigfield, 2020). In Paper 2, special changes in the learning contexts were further considered. When emergency remote teaching was put in place, students as well as teachers experienced several challenges (see, e.g., Lockl et al., 2021; Schwab et al., 2022). In this context, another advantage of the evidence-based learning activities became evident. Evidence-based learning activities can be considered a chance for university teachers and students to improve teaching and learning because they can easily be adapted to different contexts and were associated with students' motivation, even in pandemic-related distance education (Paper 2). In the case of the lecture studied, the learning activities may even have mitigated possible negative effects on student motivation in emergency distance learning.

Last but not least, Paper 3 considered the increased demand for self-regulation and its necessary prerequisites. In this study, metacognitive accuracy was added as a desirable learning outcome. We found increases in students' metacognitive accuracy in two different indicators (correct confidence and accuracy) over the course of one semester. And again, the more students used the optional activities, the greater their awareness of their own content-specific strengths and weaknesses ($\Delta R^2 = .05$) and accuracy of self-assessment ($\Delta R^2 = .04$) beyond their individual learning prerequisites and prior achievement. Other studies that investigated the effects of testing on students' correct confidence also yielded improvements with small to medium effect sizes (Barenberg & Dutke, 2019; Carvalho et al., 2022; Naujoks et al., 2022). However, most studies investigated effects of practice testing only, oftentimes neglecting the use of other learning activities (see Naujoks et al., 2022).

Bivariate associations between single learning activities and learning outcomes were inspected in an exploratory fashion. In all three studies, regularly attending classes, reviewing lecture slides and notes, participating in tests of knowledge, and submitting essays were associated with the investigated learning outcomes. Although these results were only exploratory, I consider it noteworthy that *all* of the learning activities were associated with *all* of the learning outcomes. These findings complement many studies that have largely focused on the benefits of practice testing for knowledge acquisition or metacognitive accuracy (e.g., Barenberg & Dutke, 2019; Naujoks et al., 2022; Yang et al., 2021).

7.3 Individual Learning Prerequisites Explain the Use of Learning Activities

When examining the use of learning activities, Paper 1 revealed a gap between students' intentions to use certain activities and their actual use of the activities. Considering the effectiveness of the learning activities, this gap creates the need to identify the variables associated with increased activity engagement.

In Figure 1 (see Section 2, p. 10), individual learning prerequisites are shown to shape students' learning behavior. In Paper 1, we investigated the role of motivation in students' active engagement. We found that motivational values explained variance in students' learning activity use. The more interesting, useful, and important the students' ratings of the subject, the more they engaged in learning activities in that subject. This finding is in line with previous research that investigated the mediating role of engagement between motivation and achievement (e.g., Putwain et al., 2019). Interestingly, only motivational values could explain learning behavior, but expectancies could not (cf. Putwain et al., 2019). However, in line with EVT (Eccles et al., 1983; Eccles & Wigfield, 2020), it is reasonable that the value component might explain the choice of specific learning activities when they are optional (see also Musu-Gillette et al., 2015). Teachers could address students' values in lectures by highlighting the usefulness and importance of a subject and by encouraging students to identify which aspects of the subject they find personally interesting, useful, and important (see, e.g., Yeager et al., 2014).

Although motivation could explain learning activity use in part, the rather large intention-behavior gap found in Paper 1 still needs to be discussed. Such a gap between intentions and behavior has been found in many domains of self-regulated behavior, such as health behavior (e.g., Sniehotta et al., 2005). In this dissertation's context, the gap could indicate a failure in the self-regulation of learning. It is additionally possible that students started the course with unrealistically high intentions because they did not have an accurate idea of how much time each activity would take or how much time they would need for other study related tasks. Future research could examine factors that help students set realistic intentions and put their intentions into practice. One possibility could be to make learning activities more attractive, for example, by showing students the results of studies—such as the ones presented here—that underline the usefulness of the activities or by examining barriers that hinder activity engagement. Additional ideas for promoting students' engagement in activities would be to support their self-regulated learning more generally. Some promising interventions have already been investigated, including the web-based training of self-regulated learning (Bellhäuser et al., 2016) and interventions that include implementation intentions regarding the use of optional activities (Janson & Dickhäuser, 2019; van der Beek et al., 2020).

7.4 Strengths, Limitations, and Future Research

The dissertation includes three empirical studies that longitudinally investigated the use of multiple evidence-based activities in the field. Although this kind of research came with many challenges, such as drop-out rates, potential confounding variables, and changes in the course format due to the pandemic-related lockdowns in 2020, it provides valuable insights into learning at university and has high external validity. The studies took into consideration the idea that students often have choices about how to learn, they need to maintain their effort as well as their motivation over time, and they need to accurately assess their own knowledge. The studies further investigated the same specific activities in five different cohorts and were therefore able to replicate findings on the usefulness of different activities by applying different data analytic methods. Although the activities were implemented in an introductory lecture class on educational psychology, they can easily be applied to other large lecture classes. By tackling changeable aspects of teaching in higher education and providing evidence for the usefulness of various activities, a valuable extension of the empirical basis for improvements in higher education teaching and learning has been created.

Of course, no causal conclusions about the activities' effects can be drawn because no experimental design was used. Therefore, implications for practice should be considered with caution (see also Brady et al., 2023). However, to obtain a realistic picture of students' learning behavior in a lecture class and to account for different learning activities, the studies captured students' use of optional activities rather than manipulating activity use in an experimental design. By using a longitudinal design and controlling for important predictors of learning outcomes (e.g., prior achievement), we can assume that the use of learning activities over time contributes to students' learning success.

One limitation of the longitudinal field studies was high dropout rates. The results are potentially biased, for example, because more motivated students participated in more surveys. However, in all the studies, even in the sample of students who were potentially more motivated, the use of learning activities explained learning outcomes beyond learning prerequisites and prior achievement. Further, in all studies, dropout analyses were applied, almost always indicating that the learning prerequisites of the students who participated in all the surveys did not differ from the prerequisites of those who dropped out. Only in Paper 3 did students who participated in all the surveys have better prior achievement. The differences in prior achievement may suggest that students with better learning habits (and consequently better high school grade point averages) continue their engagement, participate in studies as well as in optional learning activities, and consequently learn more. However, as this difference was found only in one of five cohorts, it is unlikely that the results were systematically biased. Further, as prior achievement was controlled for and activity use still had enough variance to explain the outcomes, prior achievement could not have been the only explanation for the learning outcomes. High dropout rates across all three Papers might be due to the voluntary participation. One possibility to diminish dropout rates would be to make participation in surveys part of the course credit. However, we decided to keep the participation voluntary to obtain high data quality and to avoid reactance on the part of the students.

For some of the analyses, the sample size was critically small. When adopting complex methods, such as in Paper 2 and Paper 3, only the minimal requirements for sample sizes were met, and the results were interpreted cautiously, especially for the latent profile analysis. If more students participated in the online surveys, the power would increase and more accurate estimates of the coefficients could potentially be obtained. However, in the field, sample size is often limited. The learning activities were implemented in a lecture class, and therefore, the sample was limited to the students attending this class. Nonetheless, future studies may benefit from implementing these activities in more and different lecture classes, thereby increasing the sample size and at the same time investigating the generalizability of the results. Whereas the focus on students attending one lecture class has the advantage of holding many relevant factors constant (e.g., content of the lecture and knowledge tests, teacher, learning activities), it also limits the generalizability of the results. Nonetheless, the results are consistent with previous research on the benefits of evidence-based learning activities for different learning outcomes (e.g., Barenberg & Dutke, 2019; Bosch et al., 2021; Dunlosky et al., 2013; Naujoks et al., 2022). Future research may explore diverse samples across different learning contexts, maybe also applying experimental designs, to underscore the significance of offering evidence-based learning activities in higher education, irrespective of the field of study.

One methodological point to consider is the assessment of participation in knowledge tests. The invitation to take part in voluntary online surveys was offered immediately after the knowledge test, likely increasing the chances of survey participation. This immediacy could have resulted in reduced variance in the knowledge test as a learning activity, potentially underestimating the association between test participation and the outcomes. Future research should aim to separate survey administration from individual learning activities and strive to objectively measure students' activity use. In the current study design, objective activity assessment was not feasible due to survey anonymity. We cannot rule out the possibility that students' self-reports of their learning activity use were inaccurate (see, e.g., Blasiman et al., 2017). However, by enhancing the specificity of the assessment from Paper 1 to Paper 3 and regularly inquiring about the number of specific opportunities students used, we likely increased students' ability to accurately report their learning activity use.

The assessment of learning activities could be improved by assessing them automatically (e.g., essay submissions) or more frequently. Beyond that, how much time students invest in each learning activity or how carefully they engage in the activities could make a difference (see, e.g., Engelschalk et al., 2017, for similar thoughts about the quality of motivation regulation strategy application). Further, how students choose their learning activities and whether this choice meets individual learning needs as well as individual goals (e.g., just passing an exam rather than getting a good grade, perhaps to have more resources to pursue other goals) could be further investigated to understand all the facets of self-regulated learning.

7.5 Implications

Keeping these limitations in mind, the studies in this dissertation revealed interesting and important results for research, theory, and practice. In future research, it would be valuable to employ designs that can simultaneously assess various learning activities, both objectively and while considering individual student preferences, needs, goals, and the ability to select suitable activities (emphasizing the need for monitoring and self-regulation). In this regard, the use of formative self-assessment, based on feedback from knowledge test and essay performance, could serve as a valuable tool for assessing students' individual learning needs. Furthermore, it would be valuable to investigate intraindividual differences in learning activity use over time, as it may exhibit substantial variability (see also Breitwieser et al., 2022). From a theoretical perspective, the results are aligned with the core assumptions of models of self-regulated learning. These include the importance of setting self-directed individual goals and planning activities (as observed in Papers 1–3, referencing Zimmerman, 2002), the necessity of sustaining motivation (as seen in Paper 2, referencing Pintrich, 2004), and the significance of accurately monitoring one's own learning progress for success (as demonstrated in Paper 3, in line with Thiede et al., 2003, and Winne, 1996). As was suggested in a supply-use model (Figure

1), several reciprocal effects were empirically supported, for example, motivation as a prerequisite for but also as an outcome of activity use (see also EVT; Eccles & Wigfield, 2020). As more proximal predictors of learning outcomes, however, specific activities were investigated in multiple cohorts in this dissertation. Based on this design and the results obtained through various statistical analysis methods, practical implications can be drawn. By examining various learning activities and their associations with multiple indicators of learning success in a real-world learning situation, we extended the empirical basis and reinforced the importance of providing evidence-based activities in higher education. This extension will enable university teachers to engage in evidence-based teaching, even in large lecture classes (see Dunn et al., 2013).

Revisiting the supply-use model, once teachers provide learning opportunities, it is essential for students to make use of these opportunities (see also Halpern & Hakel, 2003). However, students in the presented studies did not use all the learning opportunities they were offered and did so less than they had intended to at the beginning of the course. Although this result was not surprising given the multiple challenges that students face, it should nonetheless be taken seriously. If students repeatedly fail to achieve their goals, they are at risk of not completing their university courses (as observed in Benden & Lauermann, 2021) or, in the worst-case scenario, of dropping out of their study programs (as indicated by Heublein & Wolter, 2011). Therefore, researchers and practitioners alike should try to engage students in interventions that help learners initiate studying (see, e.g., Breitwieser et al., 2022).

7.6 Conclusion

This dissertation was aimed at investigating university students' use of evidence-based learning activities and at determining whether their use of such activities explains several desirable learning outcomes. In three papers, the use of specific learning activities and the associated learning outcomes were analyzed with different methodological approaches. In all samples and across all data analytic methods—hierarchical regression analyses, multigroup latent growth models, latent multivariate regression analyses, and comparisons of latent groups' outcomes—the more students engaged in optional evidence-based learning activities, the more they learned, the better they maintained their motivation over the course of one semester, and the more accurate their self-assessments of their own knowledge were at the end of the course beyond their individual learning prerequisites and prior achievement. For research and practice, the presented studies strengthen the evidence on effective learning activities in the field, allowing teachers to engage in evidence-based teaching. The presented papers offer insights to teachers on how to design their classes to enhance their students' knowledge, motivation, and metacognition. Additionally, they highlight the role students play in actively contributing to their success beyond their learning prerequisites.

References

Artelt, C. (2000). Wie prädiktiv sind retrospektive Selbstberichte über den Gebrauch von Lernstrategien für strategisches Lernen?. [How predictive are self-reported strategies for their actual use?]. *Zeitschrift für Pädagogische Psychologie, 14*(2–3), 72−84. https://doi.org/10.1024//1010-0652.14.23.72

Bae, C. L., Therriault, D. J., & Redifer, J. L. (2019). Investigating the testing effect: Retrieval as a characteristic of effective study strategies. *Learning and Instruction*, *60*, 206–214. https://doi.org/10.1016/j.learninstruc.2017.12.008

Barenberg, J., & Dutke, S. (2019). Testing and metacognition: Retrieval practise effects on metacognitive monitoring in learning from text. *Memory*, *27*(3), 269−279. https://doi.org/10.1080/09658211.2018.1506481

Bellhäuser, H., Lösch, T., Winter, C., & Schmitz, B. (2016). Applying a web-based training to foster self-regulated learning—Effects of an intervention for large numbers of participants. *The Internet and Higher Education*, *31*, 87−100. https://doi.org/10.1016/j.iheduc.2016.07.002

Benden, D.K., & Lauermann, F. (2021). Students' motivational trajectories and academic success in math-intensive study programs: Why short-term motivational assessments matter. *Journal of Educational Psychology, 114*(5), 1062. https://doi.org/10.1037/edu0000708

Bensley, D. A., Rainey, C., Lilienfeld, S. O., & Kuehne, S. (2015). What do psychology students know about what they know in psychology? *Scholarship of Teaching and Learning in Psychology*, *1*(4), 283−297. https://doi.org/10.1037/stl0000035

Bernhard, N. (2021). Students' differences, societal expectations, and the discursive construction of (de)legitimate students in Germany. *Social Inclusion*, *9*(3), 394−403.

Blasiman, R. N., Dunlosky, J., & Rawson, K. A. (2017). The what, how much, and when of study strategies: Comparing intended versus actual study behaviour. *Memory*, *25*(6), 784–792. https://doi.org/10.1080/09658211.2016.1221974

Bosch, E., Seifried, E., & Spinath, B. (2021). What successful students do: Evidence-based learning activities matter for students' performance in higher education beyond prior knowledge, motivation, and prior achievement. *Learning and Individual Differences*, *91*, 102056. https://doi.org/10.1016/j.lindif.2021.102056.

Bosch, E., & Spinath, B. (2023a). Students' motivation in an online and a face-to-face semester. A comparison of initial level, development, and use of learning activities. *Zeitschrift für Psychologie, 231*(2), 93-102. https://doi.org/10.1027/2151-2604/a000519

Bosch, E., & Spinath, B. (2023b). What evidence-based learning activities help students acquire knowledge, correct confidence in their own knowledge, and accurate self-assessment? *Learning and Individual Differences, 108*, 102374. https://doi.org/10.1016/j.lindif.2023.102374

Boser, J., Scherer, S., Kuchta, K., Wenzel, S. F. C., & Horz, H. (2017). Empirically founded teaching in psychology–An example for the combination of evidence-based teaching and the scholarship of teaching and learning. *Psychology Learning & Teaching*, *16*(2), 261–275. https://doi.org/10.1177/1475725716686452

Brady, A. C., Griffin, M. M., Lewis, A. R., Fong, C. J., & Robinson, D. H. (2023). How scientific is educational psychology research? The increasing trend of squeezing causality and recommendations from non-intervention studies. *Educational Psychology Review*, *35*(1), 37.

Breitwieser, J., Neubauer, A. B., Schmiedek, F., & Brod, G. (2022). Self-regulation prompts promote the achievement of learning goals–But only briefly: Uncovering hidden dynamics in the effects of a psychological intervention. *Learning and Instruction*, *80*, 101560. https://doi.org/10.1016/j.learninstruc.2021.101560

Brühwiler, C., & Blatchford, P. (2011). Effects of class size and adaptive teaching competency on classroom processes and academic outcome. *Learning and Instruction*, *21*(1), 95−108. https://doi.org/10.1016/j.learninstruc.2009.11.004

Busato, V. V., Prins, F. J., Elshout, J. J., & Hamaker, C. (2000). Intellectual ability, learning style, personality, achievement motivation and academic success of psychology students in higher education. *Personality and Individual Differences*, *29*(6), 1057−1068. https://doi.org/10.1016/S0191-8869(99)00253-6

Butler, A. C., Karpicke, J. D., & Roediger, H. L. III. (2007). The effect of type and timing of feedback on learning from multiple-choice tests. *Journal of Experimental Psychology: Applied, 13*(4), 273–281. https://doi.org/10.1037/1076-898X.13.4.273

Butler, A. C., & Roediger, H. L. III., (2007). Testing improves long-term retention in a simulated classroom setting. *European Journal of Cognitive Psychology*, *19*, 514–527.

Carpenter, S. K. (2012). Testing enhances the transfer of learning. *Current Directions in Psychological Science*, *21*(5), 279–283. https://doi.org/10.1177/0963721412452728

Carvalho, P. F., McLaughlin, E. A., & Koedinger, K. R. (2022). Varied practice testing is associated with better learning outcomes in self-regulated online learning. *Journal of Educational Psychology, 14*(8), 1723−1742. https://doi.org/10.1037/edu0000754

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, *132*(3), 354. https://doi.org/10.1037/0033-2909.132.3.354

Christ, A. A., Capon-Sieber, V., Grob, U., & Praetorius, A. K. (2022). Learning processes and their mediating role between teaching quality and student achievement: A systematic review. *Studies in Educational Evaluation*, *75*, 101209. https://doi.org/10.1016/j.stueduc.2022.101209

Cogliano, M., Kardash, C. M., & Bernacki, M. L. (2019). The effects of retrieval practice and prior topic knowledge on test performance and confidence judgments. *Contemporary Educational Psychology*, *56*, 117–129. https://doi.org/10.1016/j.cedpsych.2018.12.001

Credé, M., & Kuncel, N. R. (2008). Study habits, skills, and attitudes: The third pillar supporting collegiate academic performance. *Perspectives on Psychological Science*, *3*(6), 425–453. https://doi.org/10.1111/j.1745-6924.2008.00089.x

Credé, M., Roch, S. G., & Kieszczynka, U. M. (2010). Class attendance in college: A meta-analytic review of the relationship of class attendance with grades and student characteristics. *Review of Educational Research*, *8*(2), 272–295. https://doi.org/10.3102/0034654310362998

Dickhäuser, O., Dinger, F. C., Janke, S., Spinath, B., & Steinmayr, R. (2016). A prospective correlational analysis of achievement goals as mediating constructs linking distal motivational dispositions to intrinsic motivation and academic achievement. *Learning and Individual Differences*, *50*, 30–41. https://doi.org/10.1016/j.lindif.2016.06.020

Dickhäuser, O., Reinhard, MA. & Englert, C. (2011). “Of course I will ...”: The combined effect of certainty and level of expectancies on persistence and performance. *Social Psychology of Education, 14,* 519–528. https://doi.org/10.1007/s11218-011-9159-x

Dobronyi, C.R., Oreopoulos, P., & Petronijevic, U. (2019). Goal Setting, Academic Reminders, and College Success: A Large-Scale Field Experiment. *Journal of Research on Educational Effectiveness, 12,* 38–66. https://doi.org/10.1080/19345747.2018.1517849

Dochy, F., Segers, M., & Buehl, M. M. (1999). The relation between assessment practices and outcomes of studies: The case of research on prior knowledge. *Review of Educational Research*, 69(2), 145–186. https://doi.org/10.3102/00346543069002145

Downs, S. D. (2015). Testing in the college classroom: Do testing and feedback influence grades throughout an entire semester? *Scholarship of Teaching and Learning in Psychology, 1*(2), 172–181. https://doi.org/10.1037/stl0000025

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, *14*(1), 4–58. https://doi.org/10.1177/1529100612453266

Dunn, D. S., Saville, B. K., Baker, S. C., & Marek, P. (2013). Evidence-based teaching: Tools and techniques that promote learning in the psychology classroom. *Australian Journal of Psychology*, *65*(1), 5−13. https://doi.org/10.1111/ajpy.12004

Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, *12*(3), 83–87. https://doi.org/10.1111/1467-8721.01235

Dutke, S., Bakker, H. E., Papageorgi, I., & Taylor, J. (2017). PLAT 16(2) 2017: Introduction to the special issue on evidence-based teaching (EBT): Examples from learning and teaching psychology. *Psychology Learning & Teaching, 16*(2), 175–178. https://doi.org/10.1177/1475725717701209

Dutke, S., & Barenberg, J. (2015). Easy and informative: Using confidence-weighted true–false items for knowledge tests in psychology courses. *Psychology Learning & Teaching*, *14*(3), 250-259. https://doi.org/10.1177/1475725715605627

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives* (pp. 75–146). Freeman.

Eccles, J.S., & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology*, *61*, 101859. https://doi.org/10.1016/j.cedpsych.2020.101859

Eckert, C., Seifried, E., Spinath, B. (2015). Heterogenität in der Hochschule aus psychologischer Sicht: Die Rolle der studentischen Eingangsvoraussetzungen für adaptives Lehren. [Heterogeneity in higher education from a psychological perspective: The role of student entry requirements for adaptive teaching.] In K. Rheinländer (Ed.), *Ungleichheitssensible Hochschullehre* (pp.257–274). Springer. https://doi.org/10.1007/978-3-658-09477-5_14

Engelschalk, T., Steuer, G., & Dresel, M. (2017). Quantity and quality of motivational regulation among university students. *Educational Psychology*, *37*(9), 1154–1170.

European Commission, Directorate-General for Employment, Social Affairs and Inclusion. (2018). *The European qualifications framework: supporting learning, work and cross-border mobility: 10th anniversary.* Publications Office. https://data.europa.eu/doi/10.2767/385613

Fend, H. (1984). Determinanten von Schulleistungen: Wie wichtig sind Lehrer? [Determinants of student outcomes: How important are teachers?]. *Unterrichtswissenschaft Zeitschrift für Lernforschung, 12*(1), 68–86.

Foster, N. L., Was, C. A., Dunlosky, J., & Isaacson, R. M. (2017). Even after thirteen class exams, students are still overconfident: The role of memory for past exam performance in student predictions. *Metacognition and Learning, 12*, 1–19. https://doi.org/10.1007/s11409-016-9158-6.

Furnham, A., Chamorro-Premuzic, T., & McDougall, F. (2002). Personality, cognitive ability, and beliefs about intelligence as predictors of academic performance. *Learning and Individual Differences*, *14*, 47–64. doi:10.1016/j.lindif.2003.08.002

Gottfried, A. E., Fleming, J. S., & Gottfried, A. W. (2001). Continuity of academic intrinsic motivation from childhood through late adolescence: A longitudinal study. *Journal of Educational Psychology, 93*(1)*,* 3–13. https://doi.org/10.1037//0022-0663.93.1.3

Garris, C.P., & Fleck, B. (2022). Student evaluations of transitioned-online courses during the COVID-19 pandemic. *Scholarship of Teaching and Learning in Psychology*, *8*(2), 119–139. https://doi.org/10.1037/stl0000229

Halpern, D. F., & Hakel, M. D. (2003). Applying the science of learning to the university and beyond: Teaching for long-term retention and transfer. *Change: The Magazine of Higher Learning*, *35*(4), 36–41. https://doi.org/10.1080/00091380309604109

Hartwig, M. K., & Dunlosky, J. (2012). Study strategies of college students: Are self-testing and scheduling related to achievement? *Psychonomic Bulletin & Review*, *19*(1), 126–134. https://doi.org/10.3758/s13423-011-0181-y

Händel, M., & Dresel, M. (2018). Confidence in performance judgment accuracy: The unskilled and unaware effect revisited. *Metacognition and Learning*, *13*, 265–285. https://doi.org/10.1007/s11409-018-9185-6

Händel, M., Harder, B., & Dresel, M. (2020). Enhanced monitoring accuracy and test performance: Incremental effects of judgment training over and above repeated testing. *Learning and Instruction, 65*, 101245. https://doi.org/10.1016/j.learninstruc.2019.101245

Helmke, A. & Brühwiler, C. (2018). Determinanten der Schulleistung [Determinants of scholastic achievement]. In D. Rost, J. Sparfeldt, & S. Buch (Eds.), *Handwörterbuch Pädagogische Psychologie* (5th ed., pp. 78–91). Beltz.

Helmke, A. & Schrader, F.-W. (2010). Hochschuldidaktik [Higher education didactics]. In In D. Rost, J. Sparfeldt, & S. Buch (Eds.), *Handwörterbuch Pädagogische Psychologie* (4th ed., pp. 273–279). Beltz.

Helmke, A., & Weinert, F. E. (1997). Bedingungsfaktoren schulischer Leistungen. In F. E. Weinert (Ed.), *Enzyklopädie der Psychologie: Psychologie des Unterrichts und der Schule* (Vol. 3, pp. 71–176). Hogrefe.

Heublein, U., & Wolter, A. (2011). Studienabbruch in Deutschland. Definition, Häufigkeit, Ursachen, Maßnahmen [Dropout from higher education in Germany–definition, dropout rate, causes, measures]. *Zeitschrift für Pädagogik*, *57*(2), 214–236. https://doi.org/10.25656/01:8716

Hidi, S., & Renninger, K.A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*(2), 111–127. https://doi.org/10.1207/s15326985ep4102_4

Hutchings, P., Huber, M. T., & Ciccone, A. (2011). *The scholarship of teaching and learning reconsidered. Institutional integration and impact* (Vol. 21). Jossey-Bass/Wiley.

Janson, M. P., & Dickhäuser, O. (2019). Prüfungsvorbereitung an der Hochschule mittels Lernsoftware: Die Effekte von Wenn-Dann-Plänen. [Exam preparation at the university using an e-learning software: The effects of if-then-plans.] *Unterrichtswissenschaft*, *47*(3), 295–312. https://doi.org/10.1007/s42010-018-0018-8

Janson, M. P., Siebert, J., & Dickhäuser, O. (2023). Everything right or nothing wrong? Regulatory fit effects in an e-learning context. *Social Psychology of Education*, *26*(1), 107–139. https://doi.org/10.1007/s11218-022-09733-3

Jonsson, B., Wiklund-Hörnqvist, C., Stenlund, T., Andersson, M., & Nyberg, L. (2021). A learning method for all: The testing effect is independent of cognitive ability. *Journal of Educational Psychology, 113*(5), 972. https://doi.org/10.1037/edu0000627

Karpicke, J. D., Butler, A. C., & Roediger, H. L. III. (2009). Metacognitive strategies in student learning: Do students practise retrieval when they study on their own? *Memory*, *17*(4), 471–479. https://doi.org/10.1080/09658210802647009

Kobrin, J. L., Patterson, B. F., Shaw, E. J., Mattern, K. D., & Barbuti, S. M. (2008). *Validity of the SAT for predicting first-year college grade point average* [Research Report No. 2008-5]. College Board.

Kornell, N., & Bjork, R. A. (2007). The promise and perils of self-regulated study. *Psychonomic Bulletin & Review*, *14*(2), 219–224. https://doi.org/10.3758/BF03194055

Kosovich, J.J., Flake, J.K., & Hulleman, C.S. (2017). Short-term motivation trajectories: A parallel process model of expectancy-value. *Contemporary Educational Psychology, 49*, 130–139. https://doi.org/10.1016/j.cedpsych.2017.01.004

Köller, O. (2012). What works best in school? Hatties Befunde zu Effekten von Schul- und Unterrichtsvariablen auf Schulleistungen [What works best in school? Hattie's findings on effects of school and teaching variables on academic achievement]. *Psychologie in Erziehung und Unterricht*, *59*, 72–78. https:// doi.org/10.2378/peu2012.art06d

Kriegbaum, K., Becker, N., & Spinath, B. (2018). The relative importance of intelligence and motivation as predictors of school achievement: A meta-analysis. *Educational Research Review*, *25*, 120–148. https://doi.org/10.1016/j.edurev.2018.10.001

Lazowski, R. A., & Hulleman, C. S. (2016). Motivation interventions in education: A meta-analytic review. *Review of Educational Research, 86*(2), 602–640. https://doi.org/10.3102/0034654315617832.

Leeming, F. C. (2002). The exam-a-day procedure improves performance in psychology classes. *Teaching of Psychology*, *29*, 210–212.

Liborius, P., Bellhäuser, H., & Schmitz, B. (2019). What makes a good study day? An intraindividual study on university students' time investment by means of time-series analyses. *Learning and Instruction*, *60*, 310–321. https://doi.org/10.1016/j.learninstruc.2017.10.006

Lockl, K., Attig, M., Nusser, L., & Wolter, I. (2021). Cognitive and Affective-Motivational Factors as Predictors of Students' Home Learning During the School Lockdown. *Frontiers in Psychology*, *12*, 751120. https://doi.org/ 10.3389/fpsyg.2021.751120

Lounsbury, J. W., Sundstrom, E., Loveland, J. M., & Gibson, L. W. (2003). Intelligence, "Big Five" personality traits, and work drive as predictors of course grade. *Personality and individual Differences*, *35*(6), 1231–1239. https://doi.org/10.1016/S0191-8869(02)00330-6

Molenaar, I., Horvers, A., & Baker, R. S. (2019). What can moment-by-moment learning curves tell about students' self-regulated learning?. *Learning and Instruction*, 101206. https://doi.org/10.1016/j.learninstruc.2019.05.003

Musu-Gillette, L. E., Wigfield, A., Harring, J. R., & Eccles, J. S. (2015). Trajectories of change in students' self-concepts of ability and values in math and college major choice. *Educational Research and Evaluation*, *21*(4), 343–370. https://doi.org/10.1080/13803611.2015.1057161

Müller, F.H., Thomas, A.E., Carmignola, M., Dittrich, A.K., Eckes, A., Großmann, N., ... & Bieg, S. (2021). University students' basic psychological needs, motivation, and vitality before and during COVID-19: A Self-Determination Theory approach. *Frontiers in Psychology*, *12,* 775804. https://doi.org/10.3389/fpsyg.2021.775804

Naujoks, N., Harder, B., & Händel, M. (2022). Testing pays off twice: Potentials of practice tests and feedback regarding exam performance and judgment accuracy. *Metacognition and Learning*, *17*, 479–498. https://doi.org/10.1007/s11409-022-09295-x

Organisation for Economic Co-operation and Development. (2018), *Education at a Glance 2018: OECD Indicators*. OECD Publishing. https://doi.org/10.1787/eag-2018-en.

Organisation for Economic Co-operation and Development. (2019), *Education at a Glance 2019: OECD Indicators*. OECD Publishing. https://doi.org/10.1787/f8d7880d-en.

Pashler, H., Bain, P. T., Bottge, B., Koedinger, K., McDaniel, M., & Metcalf, J. (2007). *Organizing instruction and study to improve student learning*. Washington, DC: National Center for Education Research, Institute of Education Science, U.S., Department of Education.

Peverly S. T., Brobst K. E., Graham M., Shaw R. (2003). College adults are not good at self-regulation: A study on the relationship of self-regulation, note taking, and test taking. *Journal of Educational Psychology*, 95, 335–346. https://doi.org/10.1037/0022-0663.95.2.335

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review*, *16*(4), 385-407. https://doi.org/1040-726X/04/1200-0385/0

Pintrich, P. R., Smith, D. A., Garcia, 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. https://doi.org/10.1177/0013164493053003024

Putwain, D. W., Nicholson, L. J., Pekrun, R., Becker, S., & Symes, W. (2019). Expectancy of success, attainment value, engagement, and achievement: A moderated mediation analysis. *Learning and Instruction, 60,* 117–125. https://doi.org/10.1016/j.learninstruc.2018.11.005

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. https://doi.org/10.1037/a0026838

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin*, *130*(2), 261−288. https://doi.org/10.1037/0033-2909.130.2.261

Robinson, K.A., Lee, Y.K., Bovee, E.A., Perez, T., Walton, S.P., Briedis, D., & Linnenbrink-Garcia, L. (2019). Motivation in transition: Development and roles of expectancy, task values, and costs in early college engineering. *Journal of Educational Psychology*, *111*(6), 1081–102. https://doi.org/10.1037/edu0000331

Roediger, H. L. III., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, *15*(1), 20−27. https://doi.org/10.1016/j.tics.2010.09.003

Rohde, T. E., & Thompson, L. A. (2007). Predicting academic achievement with cognitive ability. *Intelligence*, *35*(1), 83−92. https://doi.org/10.1016/j.intell.2006.05.004

Rost, D.H. (2009). *Intelligenz. Fakten und Mythen* [Intelligence. Facts and myths]. Beltz.

Rowland, C. A. (2014). The effect of testing versus restudy on retention: A meta-analytic review of the testing effect. *Psychological Bulletin*, *140*(6), 1432−1463. https://doi.org/10.1037/a0037559

Rutkiene, A. & Tandzegolskiene, I. (2015). Students' attitude towards learning methods for self-sufficiency development in higher education. *Society. Integration. Education, 2015*(1), 348−357. https://doi.org/10.17770/sie2015vol1.291

Sackett, P. R., Kuncel, N. R., Arneson, J. J., Cooper, S. R., & Waters, S. D. (2009). Does socioeconomic status explain the relationship between admissions tests and post-secondary academic performance? *Psychological Bulletin*, *135*(1), 1−22. https://doi.org/0.1037/a0013978

Saville, B. K. (2010). Using evidence-based teaching methods to improve education. In S. A. Meyers & J. R. Stowell (Eds.), *Essays from excellence in teaching* (Vol. 9, pp. 48–54). Society for the Teaching of Psychology. http://teachpsych.org/ebooks/eit2009/index.php

Schindler, J., & Richter, T. (2023). Text generation benefits learning: a meta-analytic review. *Educational Psychology Review,* 35(2), 44. https://doi.org/10.1007/s10648-023-09758-w

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. https://doi.org/10.1016/j.cedpsych.2005.02.002

Schneider, M., & Preckel, F. (2017). Variables associated with achievement in higher education: A systematic review of meta-analyses. *Psychological Bulletin*, *143*(6), 565−600. https://doi.org/10.1037/bul0000098

Schwab, C., Frenzel, A. C., Daumiller, M., Dresel, M., Dickhäuser, O., Janke, S., & Marx, A. K. (2022). "I'm tired of black boxes!": A systematic comparison of faculty well-being and need satisfaction before and during the COVID-19 crisis. *PLoS ONE*, *17*(10), e0272738. https://doi.org/10.1371/journal.pone.0272738

Schwartz, B. M., & Gurung, R. A. R. (Eds.). (2012). *Evidence-based teaching for higher education*. American Psychological Association. https://doi.org/10.1037/13745-000

Schwinger, M., Steinmayr, R. & Spinath, B. (2012). Not all roads lead to Rome – Comparing different types of motivational regulation profiles. *Learning and Individual Differences*, *22*, 269–279. htttps://doi.org/10.1016/j.lindif.2011.12.006

Schwinger, M., & Stiensmeier-Pelster, J. (2012). Effects of motivational regulation on effort and achievement: A mediation model. *International Journal of Educational Research, 56,* 35–47. https:// doi.org/10.1016/j.ijer.2012.07.005.

Seidel, T. (2014). Angebots-Nutzungs-Modelle in der Unterrichtspsychologie. Integration von Struktur-und Prozessparadigma [Utilization-of-learning-opportunities models in psychology of instruction: Integration of the paradigms of structure and process]. *Zeitschrift für Pädagogik*, *60*(6), 850–866. URN: urn:nbn:de:0111-pedocs-146864

Seifried, E., Eckert, C., & Spinath, B. (2018). Optional learning opportunities: Who seizes them and what are the learning outcomes?. *Teaching of Psychology*, *45*(3), 246–250. https://doi.org/10.1177/0098628318779266

Seifried, E., & Spinath, B. (2021). Using formative self-assessment to improve teaching and learning in educational psychology courses. In S. A. Nolan, C. M. Hakala, & R. E. Landrum (Eds.), *Assessing undergraduate learning in psychology: Strategies for measuring and improving student performance* (pp. 161–176). American Psychological Association. https://doi.org/10.1037/0000183-012

Slavin, R. E. (2008). Perspectives on evidence-based research in education – what works? Issues in synthesizing educational program evaluations. *Educational Researcher, 37*(1), 5–14. https://doi.org/10.3102/0013189X08314117

Sniehotta, F. F., Scholz, U., & Schwarzer, R. (2005). Bridging the intention–behaviour gap: Planning, self-efficacy, and action control in the adoption and maintenance of physical exercise. *Psychology & Health*, *20*(2), 143–160. https://doi.org/10.1080/08870440512331317670

Spinath, B., & Seifried, E. (2012). Forschendes Lehren: Kontinuierliche Verbesserung einer Vorlesung [Research-oriented teaching: continuous improvement of a lecture.]. In M. Krämer, S. Dutke, J. Barenberg (Eds.), *Psychologiedidaktik und Evaluation IX* (pp.171−180). Shaker.

Spinath, B., & Seifried, E. (2018). Was brauchen wir, um solide empirische Erkenntnisse über gute Hochschullehre zu erhalten? [What do we need in order to obtain solid empirical findings on good university teaching?] *Zeitschrift für Hochschulentwicklung*, *13*(1), 153−169.

Spinath, B., Seifried, E., & Eckert, C. (2014). Forschendes Lehren: Ein Ansatz zur kontinuierlichen Verbesserung von Hochschullehre [Research-oriented teaching: An approach to the continuous improvement of teaching in higher education]. *Journal Hochschuldidaktik, 25*, 14−16. https://doi.org/10.17877/DE290R-6737

Steel, P., & Ferrari, J. (2013). Sex, education and procrastination: An epidemiological study of procrastinators' characteristics from a global sample. *European Journal of Personality, 27*(1), 51–58. https://doi.org/10.1002/per.1851

Thiede, K. W., Anderson, M., & Therriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology*, *95*(1), 66−73. https://doi.org/10.1037/0022-0663.95.1.66

Thomas, R. C., & McDaniel, M. A. (2013). Testing and feedback effects on front-end control over later retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(2), 437–450. https://doi.org/10.1037/a0028886

Thompson, R. A., & Zamboanga, B. L. (2004). Academic aptitude and prior knowledge as predictors of student achievement in introduction to psychology. *Journal of Educational Psychology*, *96*(4), 778−784. https://doi.org/10.1037/0022-0663.96.4.778

UNESCO. (2012). International standard classification of education (ISCED) 2011. Montreal, Canada: UNESCO Institute for Statistics.

Usher, E. L., Golding, J. M., Han, J., Griffiths, C. S., McGavran, M. B., Brown, C. S., & Sheehan, E. A. (2021). Psychology students' motivation and learning in response to the shift to remote instruction during COVID-19. *Scholarship of Teaching and Learning in Psychology*. Advance online publication. https://doi.org/10.1037/stl0000256

van der Beek, S., Bellhäuser, H., & Hertel, S. (2021). Do minimal interventions increase participation rates in voluntary online training at high school?. *Psychology Learning and Teaching*, *20*(3), 348–363. https://doi.org/10.1177/1475725720965002

Vieluf, S., Praetorius, A. K., Rakoczy, K., Kleinknecht, M., & Pietsch, M. (2020). *Angebots-Nutzungs-Modelle der Wirkweise des Unterrichts. Ein kritischer Vergleich verschiedener Modellvarianten.* [Supply-use models of the effectiveness of teaching. A critical comparison of different model variants.] In A.-K. Praetorius, J. Grünkorn, & E. Klieme, (Eds.), *Empirische Forschung zu Unterrichtsqualität. Theoretische Grundfragen und quantitative Modellierungen.* (pp. 63–80). Weinheim; Basel: Beltz Juventa

Watts, M., & Schaur, G. (2011). Teaching and assessment methods in undergraduate economics: A fourth national quinquennial survey. *The Journal of Economic Education*, *42*(3), 294–309. https://doi.org/10.3200/JECE.39.3.273-286

Wigfield, A., & Eccles, J.S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology,* *25*(1), 68–81. https://doi.org/10.1006/ceps.1999.1015

Wigfield, A., Eccles, J. S., Fredricks, J. A., Simpkins, S., Roeser, R. W. & Schiefele, U. (2015). Development of achievement motivation and engagement. In M. E. Lamb & R. M. Lerner, (Eds.), *Handbook of child psychology and developmental science: Socioemotional processes* (pp. 657–700). John Wiley & Sons.

Winne, P. H. (1996). A metacognitive view of individual differences in self-regulated learning. *Learning and Individual Differences, 8,* 327–353. https://doi.org/10.1016/S1041-6080(96)90022-9.

Wong, Z. Y., Liem, G. A. D., Chan, M., & Datu, J. A. D. (2023). Student engagement and its association with academic achievement and subjective well-being: A systematic review and meta-analysis. *Journal of Educational Psychology*. Advance online publication. https://doi.org/10.1037/edu0000833

Yang, C., Luo, L., Vadillo, M. A., Yu, R., & Shanks, D. R. (2021). Testing (quizzing) boosts classroom learning: A systematic and meta-analytic review. *Psychological Bulletin*, *147*(4), 399–435. https://doi.org/10.1037/bul0000309

Yeager, D. S., Henderson, M. D., Paunesku, D., Walton, G. M., D'Mello, S., Spitzer, B. J., & Duckworth, A. L. (2014). Boring but important: a self-transcendent purpose for learning fosters academic self-regulation. *Journal of Personality and Social Psychology*, *107*(4), 559–580. https://doi.org/10.1037/a0037637

Zimmerman, B. J. (1986). Becoming a self-regulated learner: Which are the key subprocesses?. *Contemporary Educational Psychology*, *11*(4), 307–313.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, *41*(2), 64–70.

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal*, *45*(1), 166–183. https://doi.org/10.3102/0002831207312909

Declaration in accordance to § 8 (1) c) and d) of the doctoral degree regulation of the Faculty

FAKULTÄT FÜR VERHALTENS- UND EMPIRISCHE KULTURWISSENSCHAFTEN

UNIVERSITÄT HEIDELBERG ZUKUNFT SEIT 1386

Promotionsausschuss der Fakultät für Verhaltens- und Empirische Kulturwissenschaften der Ruprecht-Karls-Universität Heidelberg / Doctoral Committee of the Faculty of Behavioural and Cultural Studies of Heidelberg University

Erklärung gemäß § 8 (1) c) der Promotionsordnung der Universität Heidelberg für die Fakultät für Verhaltens- und Empirische Kulturwissenschaften / Declaration in accordance to § 8 (1) c) of the doctoral degree regulation of Heidelberg University, Faculty of Behavioural and Cultural Studies

Ich erkläre, dass ich die vorgelegte Dissertation selbstständig angefertigt, nur die angegebenen Hilfsmittel benutzt und die Zitate gekennzeichnet habe. / I declare that I have made the submitted dissertation independently, using only the specified tools and have correctly marked all quotations.

Erklärung gemäß § 8 (1) d) der Promotionsordnung der Universität Heidelberg für die Fakultät für Verhaltens- und Empirische Kulturwissenschaften / Declaration in accordance to § 8 (1) d) of the doctoral degree regulation of Heidelberg University, Faculty of Behavioural and Cultural Studies

Ich erkläre, dass ich die vorgelegte Dissertation in dieser oder einer anderen Form nicht anderweitig als Prüfungsarbeit verwendet oder einer anderen Fakultät als Dissertation vorgelegt habe. / I declare that I did not use the submitted dissertation in this or any other form as an examination paper until now and that I did not submit it in another faculty.

Vorname Nachname / First name Family name	Eva Sophie Bosch
Datum / Date	07.02.2024
Unterschrift / Signature	

Publications of the Publication-Based Dissertation

Paper I

This is the accepted version of the article

Bosch, Eva, Seifried, Eva, & Spinath, Birgit (2021). What successful students do: Evidence-based learning activities matter for students' performance in higher education beyond prior knowledge, motivation, and prior achievement. *Learning and Individual Differences*, *91*, 102056. https://doi.org/10.1016/j.lindif.2021.102056

CRediT author statement

Eva Bosch: conceptualization, investigation, data curation, formal analysis, writing – original draft.

Eva Seifried: investigation, writing – review & editing

Birgit Spinath: conceptualization, supervision, investigation, writing – review & editing.

What Successful Students Do: Evidence-Based Learning Activities Matter for Students' Performance in Higher Education Beyond Prior Knowledge, Motivation, and Prior Achievement

Eva Bosch[1*], Eva Seifried[1] & Birgit Spinath[1]

[*]corresponding author; Eva.Bosch@psychologie.uni-heidelberg.de

[1]Heidelberg University, Department of Psychology, Hauptstraße 47-51, 69117 Heidelberg, Germany

Abstract

In higher education, students must manage their learning on their own. When students seize the opportunity to engage in specific evidence-based learning activities, this should contribute to their achievement beyond their individual learning prerequisites (i.e., prior knowledge and motivation) and their prior achievement. In turn, students with higher motivation should use more learning activities. To test these hypotheses, two cohorts of students attending a lecture class on educational psychology participated in online-surveys at the beginning and the end of one semester ($N_1 = 112$; $N_2 = 171$). Using regression analyses, we found that learning activity use explained students' performance at the end of the semester beyond their learning prerequisites and prior achievement. Furthermore, students who valued educational psychology more used more learning activities. Overall, students used learning activities much less than intended at the beginning of the semester. In conclusion, the results point to the importance of students' learning behaviors and their potential to determine their own success. Further research should identify factors that help students put their intentions into practice.

Keywords: evidence-based learning activities, motivation, higher education

1. Introduction

Previous research has already examined the roles that students' individual learning prerequisites and learning history play in their achievement in higher education: General cognitive ability, prior knowledge, and motivation are well-known individual prerequisites that explain achievement (for a meta-analysis, see Richardson et al., 2012). In addition, students experienced many different learning situations and established certain methods of learning or study habits. All these prerequisites and experiences are reflected in prior achievement, which students bring with them to higher education. When students enter higher education, they face new challenges because their learning circumstances differ from those in high school. In higher education, learning is characterized by more opportunities to choose and less external control or structure (see e.g., Morisano et al., 2010; Perry et al., 2001). For example, students often attend large lecture classes and must pass exams at the end of a semester. How they prepare for such an exam, however, completely depends on the students. These circumstances challenge students' ability to organize their learning: They alone must choose when, where, and how to study. When students are told that their prior achievement (i.e., their high school grade point average) is substantially correlated with their achievement in higher education ($r = .40$; Richardson et al., 2012), this might be demotivating for them. They could get the impression that their achievement is predetermined and consequently that their active engagement will not make a difference for their success in higher education. One aim of the present studies is to emphasize that the actual learning situation with the learning activities that are offered and how students engage in these activities contribute to students' higher education performance in ways that go beyond students' prerequisites and prior achievement.

However, it is not only students who are challenged in higher education. One aim of higher education is to equip students with competences specific for a subject but also with competences such as the ability to self-regulate their learning and thereby prepare students for lifelong learning. Therefore, instructors are challenged by the need to decide how to design their teaching to achieve these aims concurrently. Students' individual learning prerequisites become more heterogeneous as more students achieve a high school diploma that allows them to enter higher education (OECD, 2018, 2019). Students' heterogeneous learning prerequisites and the often large numbers of students attending a course challenge instructors to design their teaching in a way that enables as many students as possible to learn as much as possible. One way to address these challenges is to implement evidence-based learning activities in courses (Boser et al., 2017). Learning activities can help students learn effectively and continuously between the sessions and give instructors the opportunity to address students' knowledge level individually. Whether and to what extent students engage in such activities in turn may be explained by individual prerequisites such as motivation (Putwain et al., 2019).

This study was designed to investigate the benefits of optional evidence-based learning activities that can be offered to students in a self-regulated learning setting (e.g., a lecture class). In a lecture class, students should learn new content and can learn something about their own learning at the same time. In this setting, evidence-based learning activities serve students' learning and concurrently give instructors and researchers the opportunity to evaluate the activities' effectiveness and improve teaching. Consequently, instruction is informed by research and simultaneously, instruction aids research by supporting it with data from the field. In this field study, the use of learning activities was optional to represent a realistic learning setting in higher education. We therefore needed to assess students' intentions and actual use of learning activities before evaluating the activities' effectiveness. Three questions guided the study. First, which learning activities do students intend to use at the beginning of the semester? Second, how much do students use the learning activities that are offered over the course of the semester? Third, does the use of learning activities explain performance beyond individual learning prerequisites and prior achievement? If specific learning activities contribute to learning success in higher education beyond more stable characteristics (e.g., prior knowledge, motivation, prior achievement), this would stress the fact that students can actively contribute to their success. At the same time, it would raise the question of which factors explain students' engagement. Engagement could be explained by looking at individual learning prerequisites, again. This was considered in the second study.

2. Supply-use model for instruction

To investigate specific learning activities that can be implemented in higher education and determine whether they benefit students' learning, it is necessary to understand the general framework within which learning takes place and the factors that influence the learning process. Several authors have presented models of instruction for different contexts as an interaction of supply and use (Brühwiler & Blatchford, 2011; Helmke, 2017; Seidel, 2014). The basic idea behind these models is that learning opportunities are provided, and learners can use them to acquire knowledge. One advantage of these models is that they look at predictors of learning outcomes at several levels and from several perspectives as well as at their interactions: Characteristics of learners, instructors, and the context can be used to explain the supply and use of learning opportunities, which in turn contribute to achievement.

Figure 1 depicts a supply-use model adapted for the present study. Students' individual learning prerequisites are important in several ways. First, individual prerequisites are thought to explain achievement directly. Researchers have emphasized the roles of cognitive ability (Kuncel & Hezlett, 2007), prior knowledge (Thompson & Zamboanga, 2004), motivation (e.g., Robbins et al., 2004), and study habits (Credé & Kuncel, 2008) for achievement in higher education. Furthermore, these individual prerequisites have already contributed to prior achievement (e.g., in high school), which is also highly correlated with subsequent achievement (Schneider & Preckel, 2017).

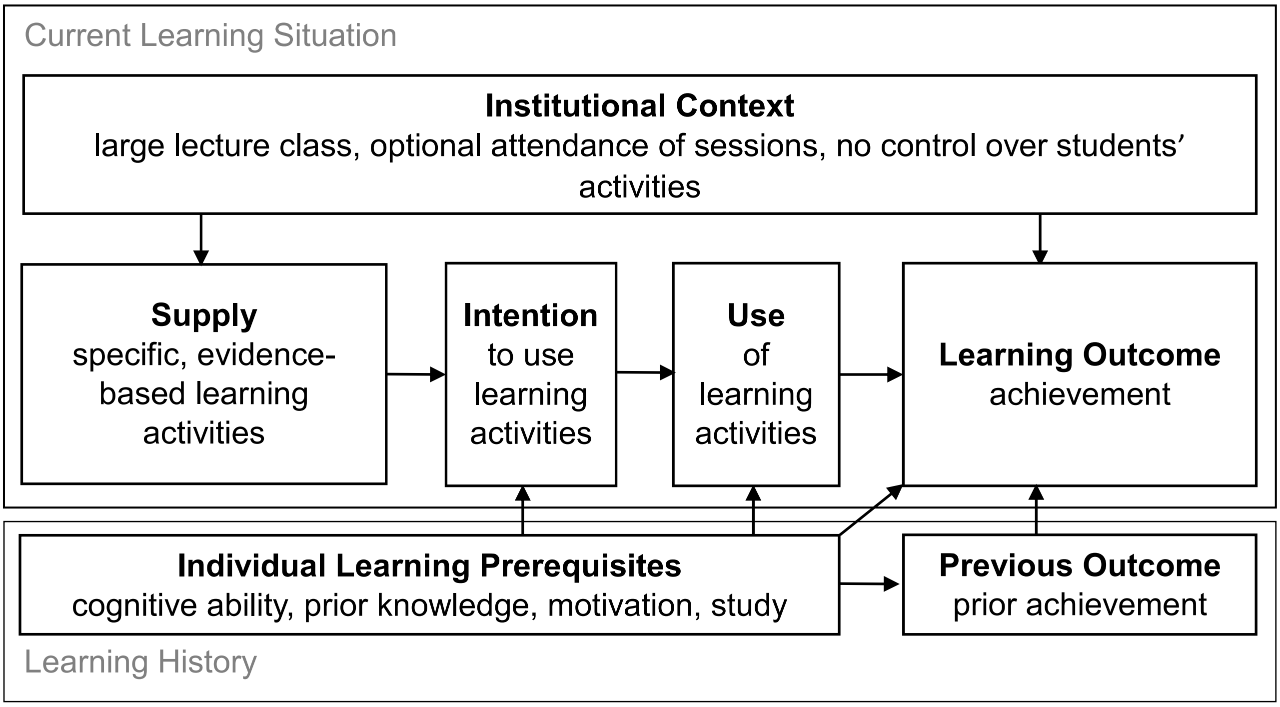


Fig. 1. A Supply-Use Model of Learning, Adapted for Higher Education

However, individual prerequisites such as motivation also influence how much students engage in their current learning situation, and this, in turn, explains achievement as well (Putwain et al., 2019). In the supply-use model (Figure 1), engagement is specified as the intention to use and the actual use of specific learning activities. Thus, we consider it important to focus on specific behavioral aspects of learning because they are more proximal antecedents than individual learning prerequisites, while of course the individual prerequisites still matter.

The specific learning activities that can and should be offered depend in part on the institutional context. For example, the institution often limits the number of students attending a course and the resources available for the course. Consequently, many students attend a lecture class for one semester, but the instructor usually does not monitor their learning progress until the end of the semester, when students must pass an exam. One opportunity for instructors to shape the learning process within this framework is to offer optional learning activities over the course of the semester. Hereby, instructors should consider evidence-based activities that rely on empirical findings regarding their effectiveness (Boser et al., 2017). To address the benefits of continuous learning (contrary to cramming at the end of the semester, only; see e.g., Dunlosky et al., 2013), instructors should offer activities regularly over the course of the semester. Regarding the specific activities, instructors should take into account beneficial effects such as testing (Carpenter, 2012), peer interaction (Schneider & Preckel, 2017), and feedback (Downs, 2015).

2.1 Individual learning prerequisites

Learners' individual learning prerequisites are well-known predictors of achievement in several learning contexts, explaining approximately 50 % of the variance in educational achievement (Köller, 2012). In a comprehensive review of meta-analyses, Schneider and Preckel (2017) summarized several categories of students' variables that are associated with achievement in higher education. For example, prior knowledge has been found to be associated with better achievement (Dochy et al., 1999; Thompson & Zamboanga, 2004). It is reasonable that students with high prior knowledge perform better on exams because they can build on their existing knowledge. Furthermore, prior knowledge in part results from prior effort, which is quite stable over time (Gottfried et al., 2001). In an introductory course in higher education, prior knowledge should also show more variance (e.g., in school) and should therefore be considered when explaining achievement. The importance of this predictor is reflected in the large effect sizes of the relationship between prior knowledge and achievement in higher education (d = 0.79; Schneider & Preckel, 2017).

Another very important predictor that Schneider and Preckel (2017) classified as a student variable is motivation. Expectancy value theory (Eccles et al., 1983; Wigfield & Eccles, 2000) implies that expectancies (how well students think they will perform) and values (how important, enjoyable, and useful a subject is for them) explain students' achievement-related choices and behavior, which in turn explain performance. These assumptions have been confirmed multiple times empirically: Expectancies along with the interaction of expectancies and values explained engagement, which in turn explained performance (e.g., Putwain et al., 2019). When explaining performance, the expectancy-related component of motivation is especially important to consider (d = 1.81 for performance self-efficacy; Schneider & Preckel, 2017), whereas values explain achievement-related choices (Musu-Gillette et al., 2015).

Prior knowledge and motivation are important learning prerequisites that students bring with them to higher education. These individual characteristics as well as experiences with learning in the past influence prior achievement. As such, prior achievement is a condensed indicator of all determinants that have been important for achievement prior to the actual learning situation, and it is a very good predictor of future achievement. In the review by Schneider and Preckel (2017), prior achievement showed (along with intelligence) the strongest relation with achievement. Especially grade point average in high school (HSGPA) has been found to be strongly associated with subsequent academic achievement (e.g., $\rho = .40$; Richardson et al., 2012; $d = 0.90$ Schneider & Preckel, 2017). Therefore, it is particularly important to show students that the use of learning activities contributes to the explanation of performance beyond prior achievement.

Students' actual learning behavior should be important for their achievement beyond individual prerequisites and prior achievement because behavior in the current learning situation is considered a more proximal predictor of achievement than individual learning prerequisites and prior achievement (see Figure 1). Because of the specific challenges of learning and instruction in higher education, however, we need to look more closely at these learning behaviors and how they can be investigated.

2.2 Institutional context: self-regulated learning in higher education

In the supply-use model (see Figure 1), learning is described as an interaction of the supply and use of learning opportunities. Whereas this general mechanism is comparable to high school (see, e.g., Brühwiler & Blatchford, 2011), higher education students have greater responsibility for their use of learning opportunities. Students need to decide on their own when, where, and how to study without external control from instructors. This has been referred to as self-regulated learning. Self-regulated learning has been defined as "the degree to which students are metacognitively, motivationally, and behaviorally active participants in their own learning process" (Zimmerman, 2008, p. 167). Students need to become self-regulated learners because, in contrast to high school, higher education provides students with more opportunities to organize learning individually. Students experience less external control and therefore need to take responsibility and actively engage in their own progress (Larose et al., 2005).

Within different theoretical frameworks of self-regulated learning, there are several strategies that are believed to enable students to control and direct their learning successfully in the absence of external control and are therefore considered to contribute to performance (Pintrich, 2004; Winne & Hadwin, 1998; Zimmermann, 2000; for an overview, see Boekarts et al., 2000). Despite the theoretical consensus on the importance of self-regulated learning in higher education, it is difficult to evaluate the impact of self-regulated learning on performance from empirical data. Whereas some authors have emphasized the importance of self-regulated learning strategies on the basis of empirical findings (e.g., Schneider & Preckel, 2017), others have found that such strategies do not predict performance (e.g., Richardson et al., 2012). At first glance, this may sound contradictory. However, a closer look at the constructs that are studied under the term *self-regulated learning strategies* reveals that considerable differences exist.

2.3 Approaches to self-regulated learning: learning strategies and learning activities

To facilitate a precise understanding of self-regulated learning and its contribution to students' success in higher education, we consider it useful to distinguish between (self-regulated) learning strategies and learning activities. Learning strategies have been defined as metacognitive, motivational, and behavioral strategies that enable learners to monitor and regulate their own behavior in learning situations in a goal-directed way in general (e.g., Pintrich, 2004). These strategies are often assessed with questionnaires, for example, the *Motivated Strategies for Learning Questionnaire* (MSLQ; Pintrich et al., 1993) and the

Learning and Study Strategies Inventory (LASSI; Weinstein et al., 1987). These questionnaires target the aforementioned aspects of self-regulated learning: motivation (e.g., value, expectancy, and affect), metacognition (e.g., rehearsal, elaboration, organization, critical thinking, and metacognitive self-regulation), and behavioral aspects (e.g., time and study environment regulation, effort regulation, peer learning, help seeking, and time management). Credé and Phillips (2011) meta-analyzed the association of strategies assessed with the *MSLQ* and GPA in college. They reported rather small correlation coefficients between the scales for cognitive and meta-cognitive strategies as well as effort regulation and time/study environment with GPA in college (ρ = -.11 to .23). However, when the *MSLQ* or the *LASSI* were used, no self-regulated learning strategy explained performance except for effort regulation when prior achievement was controlled for in another meta-analysis (Richardson et al., 2012). Kim et al. (2020) analyzed the associations between different areas of self-regulation and achievement and found that regulation of behavior and motivation affected achievement, whereas regulation of cognition did not. Additionally, interventions in self-regulated learning improve achievement only partially through higher self-regulated learning. Other factors might contribute to the improvement such as time on task or motivation (Jansen et al., 2019). From these results, we conclude that the motivational and behavioral levels of self-regulated learning add to the explanation of performance beyond students' individual learning prerequisites.

Targeting students' learning behavior seems a promising venue for understanding self-regulated learning better. In the literature, several promising methods have been investigated in order to assess learners' study time precisely (Liborius et al., 2019; Molenaar et al., 2021) or interventions to help learners initiate studying (Breitwieser et al., 2020). To investigate more accurately, what exactly students do when they invest time in studying seems a good addition to these studies. In this regard, students' learning behavior should be assessed by targeting *specific* activities in a *specific* learning situation. We consider such behavior *learning activities*. Learning activities comprise behaviors that students engage in to acquire knowledge. Learning activities in general have already been found to contribute to students' performance. For example, in their meta-analysis, Credé and Kuncel (2008) found that an aggregate measure of study habits (study routines, e.g., frequency of study sessions, self-testing, review of material; i.e., learning activities) explained variance in achievement in college after controlling for prior achievement ($\Delta R^2 = .03$). It is plausible that study habits contribute to students' success because "[s]tudents must act to acquire knowledge (study, practice, integrate, retain) before it can be translated into performance on a test or exam" (Credé & Kuncel, 2008, p. 430).

Consequently, we studied self-regulated learning as follows. We studied learning activities in a specific learning situation and evaluated which activities students used and how this contributed to their performance in that specific learning situation.

2.4 Supply: evidence-based learning activities in a lecture class

Students need to direct and initiate learning on their own in higher education. However, students do not always use the most effective learning activities when they need to organize their learning on their own (see, e.g., Karpicke et al., 2009; Kornell & Bjork, 2007). Instructors, however, can offer several opportunities for dealing with the to-be-learned material. When instructors make decisions about such learning opportunities, they should consider evidence-based activities. With these activities, students' learning can be optimized (Dunn et al., 2013; Pashler et al., 2007; see also recommendations by Boser et al., 2017). Several learning activities have been shown to contribute to performance, especially activities that help students space out their learning (Dunlosky et al., 2013) and test themselves (Bae et al., 2019; Cogliano, et al., 2019; Dunlosky et al., 2013; Jonsson et al., 2020; Roediger & Butler, 2011; Rowland, 2014). Testing is thought to improve students' learning because it enhances memory not only for the tested content but also for new material (Carpenter, 2012). More specifically, attending classes regularly (Credé et al., 2010), taking practice tests (see, e.g., Bae et al., 2019; Dunlosky et al., 2013), and engaging in peer interactions (Schneider & Preckel, 2017) have proven helpful for students to improve their learning outcomes. Study groups can benefit students' learning as they help students to estimate realistically how much they (do not) know in comparison with their peers and what the instructor may expect in a lecture. In a lecture class, instructors can offer such activities and explain the usefulness of these activities to help students form corresponding learning intentions, which are an important predictor of actual behavior (Gollwitzer, 1996). If students engage in these activities, performance feedback can be useful because feedback helps students see what they do and do not know and therefore what they should focus on (for benefits of feedback, see, e.g., Downs, 2015; Hattie, 2011).

Previous research has examined the relationship between academic performance and single activities (e.g., Credé et al., 2010) or the benefits of specific activities in comparison to other activities in the laboratory (e.g., testing compared to restudying, see Carpenter et al., 2008). Other researchers designed teaching including evidence-based activities but did not monitor how much students use the activities or how the use relates to performance (Boser et al., 2017). We build on these studies that investigated the effectiveness of learning activities but assess more learning activities all together in one lecture class in the field and evaluate how much students use them, when they are optional. Another, maybe more questionable learning activity, is memorizing content ahead of the exam. Although in previous studies, rereading and thereby trying to memorize is not recommended (e.g., Geller et al., 2018, Karpicke et al., 2009), studies also found that students nonetheless do so (Blasiman et al., 2017). We consequently wanted to represent this in our study next to the other activities.

Psychology instructors have the opportunity and expertise to evaluate the usefulness of their teaching, including the activities they offer (Spinath et al., 2014). Following these ideas, we designed a lecture class with evidence-based activities (all activities and their implementation in a lecture class are described in further detail by Seifried & Spinath, 2020). In the aforementioned publication the lecture design and learning activities are described in detail. However, no empirical data were collected or presented. Therefore, there is no overlap regarding data or samples. In the next step, we implemented these activities and evaluated their effectiveness by monitoring students' use of the activities and assessing their achievement. The results are presented in this article.

In sum, we explored students' intentions to use such learning activities at the beginning of a semester, to what extent they subsequently used the activities over the course of the semester, and how this use was related to performance. The literature on self-regulated learning will benefit from our study for three reasons. First, we took into account that students need to maintain their effort over the course of a whole semester without external control and captured the challenges of students' learning. Second, we investigated specific activities that target specific behavior and can therefore be assessed more accurately than learning strategies. Third, these activities can be applied even in large lecture classes to equip both instructors and students with valuable insights into the improvement of students' learning.

Thus, in Study 1, we explored the link between learning activities and achievement. If learning activities prove useful for predicting achievement, it would be worthwhile in another step to look at factors that predict the use of these learning activities. In this regard, motivation should be of specific importance. There is empirical evidence that engagement mediates the effects of motivational constructs on performance (Putwain et al., 2019). Thus, in Study 2, we hypothesized that motivation at the beginning of the semester would predict the use of specific learning activities over the course of the semester.

3. Summary and research questions

On the basis of supply-use models of learning, we investigated the use of learning activities as a predictor of achievement in higher education (see Figure 1). Several individual student characteristics (i.e., prior knowledge, motivation, prior achievement) can be expected to be related to performance in higher education (see also Richardson et al., 2012). Beyond this, however, self-regulated learning plays a crucial role in higher education (Larose et al., 2005). Questionnaires assessing self-regulatory learning strategies have mostly failed to explain students' performance beyond prior knowledge, motivation, and prior achievement (Richardson et al., 2012). However, the assessment of *specific* learning activities in a *specific* learning situation should prove more helpful. Evidence-based learning activities (see, e.g., Boser et al., 2017, Dunlosky et al., 2013; Pashler et al., 2007)—or more specifically: whether they are used—should contribute to students' learning success. However, to examine whether these activities are useful, it is necessary to control for other well-established powerful predictors of achievement. Only if the activities explain unique variance in students' achievement (as posited by theories on self-regulated learning) can they be considered useful (LeBreton et al., 2007). Because it is still unclear to what extent students implement learning activities during a semester without external control and how this relates to their performance, we posed the following research questions and hypotheses in Study 1:

RQ1) Which learning activities do students intend to use in a lecture class?
RQ2) To what extent do students actually use learning activities over the course of a semester? What are the activities that students use the most, and do students use the activities as intended at the beginning of the semester?
RQ3) To what extent is each learning activity associated with performance at the end of the semester?
H1) The use of learning activities during a semester will predict performance on a mock exam at the end of the semester beyond the individual learning prerequisites of prior knowledge and motivation.
H2) Learning activities will predict performance on a mock exam even after additionally controlling for prior achievement.

In a second study, we wanted to replicate and extend these findings by posing a hypothesis about the explanation of learning behavior:

H3) Motivation at the beginning of the semester will predict the use of learning activities over the course of the semester.

4. Study 1

4.1 Procedure

The setting of the study was an introductory lecture class on educational psychology for preservice teachers and psychology undergraduates at a German university. The aim of the lecture class was to provide students with basic knowledge regarding important educational psychological topics (e.g., how to improve learning). Furthermore, as the program came rather early in the curriculum, it was also dedicated to providing students with basic statistical knowledge to enable them to interpret empirical evidence regarding educational psychological topics (e.g., to understand the concept of correlations and the importance of experiments for causal conclusions). The lecture class included weekly lecture sessions. Besides these sessions, we offered the following learning activities with regard to evidence-based principles: Students had the opportunity to take part in practice tests five times and to submit short essays or self-generated exam questions seven times to practice for the exam. Students were given individual feedback on their submissions and practice test results so that they were able to monitor their learning progress throughout the course.

On each practice test, students were quizzed on the lectures' topics of the last three weeks online and were given feedback on their performance. During the practice phases, students were given the opportunity to answer essay questions and submit them. Essay questions were presented at the end of each lecture and required students to think about the content of the lecture and apply them to a new problem. The questions were comparable to the ones asked on the exam. Thereby, the students practiced the tasks they later faced on the exam. Furthermore, they received feedback on their submissions and could thereby become aware of any content they had not understood correctly. This task has already been shown to be beneficial for students (Seifried et al., 2018). As an alternative to the essays, students had the opportunity to submit questions that could be used for the exam and receive feedback on these questions. This task required the students to narrow down the content to the most important and specific information. The design of the specific learning activities is also described in further detail by Seifried and Spinath (2020). The practice tests and practice phases were offered several times during the semester and thereby helped the students learn continuously (i.e., space their learning). Even though attendance was not obligatory, we still recommended that they attend the lecture sessions regularly because of the positive link between frequency of class attendance and performance in earlier studies (Credé et al., 2010). There is also evidence that regularly reviewing materials (Hartwig & Dunlosky, 2012) and social interaction such as explaining the course material to others and learning together in small groups is associated with students' performance (Schneider & Preckel, 2017). Thus, we recommended these activities as well.

However, as noted above, students tend to have trouble self-regulating their engagement in such activities. Because all activities were optional for students, we explained these activities and explicitly suggested that students use them during the semester. Over the course of the semester, the (non-)use of activities was not monitored. However, we helped students form corresponding learning intentions, which could help them engage in the activities.

As part of their coursework, we asked students to complete online surveys in the first week of their semester (T1), and from this point on, approximately every 3 weeks. The last survey took place 12 weeks later, that is, 2 weeks ahead of the final exam (T2). At T1, we assessed students' prior achievement, motivation, and prior knowledge regarding educational psychology as well as their intentions to implement the suggested evidence-based learning activities during the semester. At T2, we assessed students' use of the learning activities over the whole semester, their motivation, and their performance on a mock exam. Besides the measures used for the current study, we assessed further variables within the online surveys to illustrate specific contents of the lecture. Therefore, we assessed motivational competences

(Schaller & Spinath, 2017), interest in different aspects of teaching (Mayr, 1998) and facets of teacher personality (Brandstätter & Mayr, 1994) and showed students a scale used in the diagnostic process of ADHD (Brühl et al., 2000) in study 2. At both time points, we also assessed further scales that were mostly related to the evaluation of the lecture and thus were not used in the present study. Participation in the surveys was voluntary and anonymous (students generated anonymous codes for longitudinal monitoring). The procedure was in accordance with human subjects principles and procedures in Germany, which do not require formal review for anonymous survey studies like the one presented here by default. No harm for participants was expected because they only filled in online questionnaires regarding their learning prerequisites and current knowledge. Further, students were informed that they would have no disadvantage if they did not participate. Each time, students were informed about the content and aims of the survey, and gave their consent.

4.2 Sample

At the beginning of the semester, $N = 223$ preservice teachers took part in an online survey. Of those, $n = 112$ students took part in the survey at T2 and were therefore included in the analyses. Participants were on average $M = 21.89$ years old ($SD = 3.23$), and 67.9 % were female (30.4 % male; 1.8 % did not indicate their gender). They stated that they had been studying on average for $M = 4.4$ semesters ($SD = 2.51$) and indicated studying a wide variety of teaching subjects. Missing data analyses revealed that students who continued to participate in our study had a better HSGPA ($M = 1.98$, $SD = .58$) than students who did not participate at T2 ($M = 2.24$, $SD = .59$), $t(221) = 3.34$, $p < .01$, $d = 0.44$. We computed equivalence tests, setting the equivalence bounds at $\Delta = 0.40$ because with at least 111 subjects in each condition (included and drop out) an effect of $d = 0.33$ can be detected with a probability of 80 % (G*Power, Faul et al., 2007) and is therefore considered the smallest effect size of interest here (see Lakens, 2017; Lakens et al., 2018). Equivalence tests indicated that on all other study variables, there were no further differences between students who participated at T2 and those who did not (all $p < .05$) except for values where the effect is statistically not different from zero ($t(216) = -1.89$, $p = .06$) but statistically not equivalent to zero ($t(216) = 1.05$, $p = .15$). Considering that 181 students completed the course and took part in the final exam, about 62 % of all students attending the lecture class participated in our study.

4.3 Measures

Besides the variables used for illustrating the lectures' content (see 4.1), we only assessed the following variables:

To assess *prior achievement* (T1), we asked students to indicate their HSGPA. In Germany, the HSGPA ranges from 1 to 4, with lower values representing better grades. For the analyses in this paper, we recoded the HSGPA so that higher values represented better grades.

Prior knowledge (T1) regarding educational psychology was assessed with a short test of 15 items. The items covered content that was addressed later in the lecture class. Each item consisted of a sentence that was either right or wrong (e.g., "Dyslexia usually comes along with a lower IQ"). Students had to indicate whether they thought the statement was true or false and indicated their confidence in their answer (confidence-weighted true-false items; Dutke & Barenberg, 2015). If their answer was correct, students received 1 point (irrespective of their confidence). Thus, possible scores ranged from 0 to 15.

Motivation for educational psychology (T1 and T2) was assessed with three items measuring *ability self-concept* (adapted for the "educational psychology" context instead of "school" from the Scales for measuring academic ability self-concept; SESSKO; Schöne et al., 2002) and nine items measuring *subjective values* (adapted from the Scale for Assessing Subjective School-related Values; SESSW; Steinmayr & Spinath, 2010).

Intentions and use of learning activities (T1 and T2) were assessed with eight items, each concerning a learning activity that could be used during the semester: regularly attending the lecture sessions, reviewing lecture notes or literature, submitting essays or items, taking practice tests, forming a study group, and memorizing the session's content for the exam. At

T1, students were asked to indicate how much they intended to apply each activity during the semester on a scale ranging from 1 (*not at all*) to 5 (*very much*). At T2, students were again presented with the activities and were asked how much they had used each strategy. Again, students could indicate 1 (*not at all*) to 5 (*very much*) for each strategy. We computed means across all activities for intentions at T1, use of activities at T2, and differences between the two (use (T2)-intention (T1)). In this study, the use of learning activities can be considered a formative construct (see Coltman et al., 2008). The use of very different learning activities can be viewed as diverse facets of students' engagement. Hence, students' engagement is not considered a latent construct that exists independent of the measures used but the learning activities are a set of activities that make up students' engagement. There are no assumptions about intercorrelations of the items (more use of one activity is not systematically associated with more or less use of another activity).

Performance (T2) was assessed with a mock exam consisting of 39 items that addressed content from the whole lecture class. The items had the same format as the short test for prior knowledge and the final exam. Each item consisted of one statement for which students had to indicate whether it was true or false while indicating their confidence in their answer. Scores for achievement could range from 0 to 39. In order to make results comparable between samples we performed a z-transformation.

4.4 Results

We used R (version 4.0.4, R Core Team), the TOSTER package (Lakens, 2017) for equivalence tests and IBM SPSS Statistics (version 24) for all further statistical analyses. Descriptive statistics and correlations of all study variables are presented in Tables 1 and 2. Students began the semester with high motivation and high intentions to use all learning activities. The average HSGPA found in the sample represents a *good* performance according to the German grading system. The average score for prior knowledge indicates that students began the lecture with little knowledge (because the mean score represented the probability of guessing).

Table 1

Study 1: Means and standard deviations for learning activities.

(N = 112)	Intention (T1) M (SD)	Use (T2) M (SD)	Difference Use-Intention M (SD)
Memorizing content for the exam	4.83 (.46)	4.26 (1.08)	–0.57 (1.15)
Attending lectures regularly	4.67 (.66)	4.14 (1.24)	–0.53 (1.16)
Taking practice tests regularly	4.60 (.61)	4.43 (.88)	–0.17 (1.02)
Reviewing lectures notes regularly	4.24 (.84)	3.62 (1.28)	–0.62 (1.30)
Submitting essays regularly	3.69 (.88)	2.28 (1.42)	–1.41 (1.41)
Forming a study group	3.45 (1.19)	2.72 (1.16)	–0.73 (1.31)
Submitting items regularly	3.38 (1.05)	1.73 (1.17)	–1.67 (1.50)
Reviewing literature regularly	2.58 (1.01)	1.43 (.82)	–1.16 (1.14)

Note. Intentions to use and use of learning activities could be indicated on a scale ranging from 1 (*not at all*) to 5 (*very much*).

4.4.1 Intentions and use of learning activities

The activities that students intended to use the most (see RQ1) were to memorize content for the exam (M = 4.83, SD = 0.46) and to attend the lecture regularly (M = 4.67, SD = 0.66). A paired t test with an adjusted $\alpha < .007$ (Bonferroni-corrected) indicated that memorizing content for the exam was significantly preferred over all other activities, $t(111) = 7.29$ to 21.27, all ps < .007, except attending the lecture sessions regularly, $t(111) = 2.21$, p = .028. However, students intended to use many activities a lot (almost all intentions were above the scale mean).

To address RQ2 about the actual use of learning activities during the semester, we inspected the means and standard deviations of each learning activity. The activity used the most—on a descriptive level—was practice tests (M = 4.43, SD = 0.88). However, we interpret this finding cautiously: If students had not taken the first and last practice test, we would not have their data, and they would not be included in the sample. Apart from the practice tests, students reported memorizing the content for the exam above all other activities, $t(104) = 4.34–22.23$, $p < .007$, except attending the lecture sessions regularly, $t(104) = 0.58$, $p = .56$, and taking practice tests, $t(104) = -1.38$, $p = .17$. Table 1 shows the actual use of all activities. Looking at the corresponding intentions to use each learning activity, we see that students used each activity less than intended (see Table 1). At the end of the semester, students reported that overall, they implemented their intentions less than they had intended, $t(105) = 14.45$, $p < .001$, $d = 1.61$. However, the more students intended to use the learning activities, the more they implemented the activities ($r = .34$, $p < .01$).

4.4.2 Predicting performance

Bivariate correlations (see Table 2) indicated that performance on the mock exam was related to HSGPA, ability self-concept at T2, and use of learning activities. However, prior knowledge and ability self-concept at T1 were not related to performance at the end of the semester.

Table 2

Study 1: Means, standard deviations, reliabilities, and bivariate correlations of study variables

(N = 112)	*M*	*SD*	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) HSGPA	3.02	(0.58)	-							
(2) Prior knowledge (T1)	7.58	(1.91)	–0.10	-						
(3) Intentions (T1)	3.92	(0.46)	–0.12	–0.10	-					
(4) Values (T1)	3.93	(0.69)	0.06	–0.12	0.49**	(.89)				
(5) Self-concept (T1)	3.47	(0.66)	0.16	–0.02	–0.04	0.14	(.85)			
(6) Values (T2)	3.69	(0.74)	0.06	–0.10	0.24*	0.61**	0.21*	(.92)		
(7) Self-concept (T2)	3.15	(0.74)	0.12	–0.06	–0.02	0.10	0.35**	0.50**	(.80)	
(8) Learning activities (T2)	3.08	(0.58)	0.03	0.01	0.34**	0.28**	0.01	0.26**	0.07	-
(9) Performance (T2)	28.73	(4.04)	0.35**	0.00	–0.03	0.02	0.06	0.10	0.25*	0.23*

Note. Self-report measures all used a response scale from 1 to 5, except for high school grade point average (HSGPA; possible range: 1 to 4), grades were recoded: larger numbers represented better grades. Prior knowledge could range from 0 to 15; performance could range from 0 to 39. Cronbach's alpha, when applicable, is listed on the diagonal.
* p < .05. ** p < .01.

To explain performance at T2, we computed hierarchical regressions (H1 and H2). Because prior knowledge, subjective values, and ability self-concept at T1 were not associated with performance at T2, we excluded these constructs from further analyses and the regression model. In a first step, we entered significant correlates of performance at T2, namely, self-concept at T2 and the learning activities as predictors in the analysis. In a second step, we entered prior achievement into the model. Results are depicted in Table 3.

Table 3

Study 1: Hierarchical regression performance on the mock exam (z-transformed)

Variable	*B*	95 % CI for *B*		*SE B*	ß	R^2	ΔR^2
		LL	*UL*				
Step 1						0.08	0.08**
Constant	–2.03**	–3.25	–0.82	0.61			
Self-concept (T2)	0.29*	0.03	0.55	0.13	0.21*		
Learning activities (T2)	0.36*	0.02	0.69	0.17	0.20*		
Step 2						0.17	0.09**
Constant	–3.46	–4.89	–2.02	0.72			
Self-concept (T2)	0.25	0.00	0.50	0.13	0.18		
Learning activities (T2)	0.35*	0.03	0.67	0.16	0.20*		
HSGPA	0.53**	0.22	0.84	0.16	0.30*		

Note. N = 112. CI = confidence interval, HSGPA = high school grade point average.
* p < .05. ** p < .01.

Learning activities explained performance on the mock exam (ß = 0.20, p < .05) after controlling for ability self-concept (ß = 0.21, p < .05). This held after entering prior achievement into the model (ß = 0.20, p < .05, for learning activities; ß = 0.30, p < .01, for prior achievement). In the second model with prior achievement as a predictor, ability self-concept no longer predicted performance on the mock exam (ß = 0.18, $p = .05$). In the final model, all predictors together accounted for 17 % of the variance in performance.

Goodman and Kruskal's Gamma correlations were computed to investigate bivariate associations between each of the eight learning activities and performance on the mock exam. Results are depicted in Table 4.

Table 4

Study 1: Bivariate associations of single learning activities and performance on the mock exam.

($N = 112$)	Memorizing content for the exam	Attending lectures regularly	Taking practice tests regularly	Reviewing lecture notes regularly	Submitting essays regularly	Forming a study group	Submitting items regularly	Reviewing literature regularly
Performance (T2)	0.24*	0.14	0.02	0.21*	0.20*	–0.05	0.04	–0.01

Note. Use of learning activities could be indicated on a scale ranging from 1 (*not at all*) to 5 (*very much*). Performance could range from 0 to 39. * $p < .05$.

4.5 Discussion

In this study, we offered several evidence-based activities to the students in a lecture class. At the beginning of the semester, students strongly intended to use every activity (RQ1). At the end of the semester, however, they indicated that they had used them much less than intended. On average, every second activity was used *rather not* or *not at all*. The activities used the most were memorizing content for the exam, attending lectures regularly, and taking practice tests regularly (RQ2). This illustrates the typical challenge with self-regulated learning: Students have trouble maintaining their effort and engagement over the course of a semester in a self-regulated learning setting. We found a gap between intentions and behavior, a gap that has already been found in other disciplines (e.g., physical activity; Sniehotta et al., 2005). In this regard, we looked at predictors for learning activity use in a next step in Study 2.

When students used the learning activities, this positively contributed to their performance at the end of the semester (H1). This was still the case after controlling for students' prior achievement (H2). Thus, the activities were useful for our students in that the activities made a difference for students' achievement. The investigation of specific activities in turn provides ideas about which activities could be implemented even in large lecture classes to support students' success.

5. Study 2

Motivation is an important predictor for achievement (see also 2.1 Learning Prerequisites). In Study 2, this assumption was specified and investigated: motivation is thought to drive behavior and expectancies and values as motivational constructs are associated with achievement related behavior and choices (see e.g., Wigfield & Eccles, 2000). These motivational constructs should consequentially also explain learning behavior or, in this case, the use of specific learning activities in higher education. Study 2 was designed to replicate and expand the findings from Study 1. To this end, we investigated the aforementioned research questions and hypotheses in another sample and additionally examined the role of motivation in explaining the use of learning activities.

5.1 Procedure

The setting for Study 2 was the same as in Study 1 (i.e., an obligatory lecture on educational psychology), but the lecture class took place one semester later. Regarding the contents of the lecture there were certain differences, however, all learning activities, their frequency and requirements for students were the same as in study 1. In the first session, we asked students to complete an online survey in the first week of the semester (T1) and afterwards approximately every 3 weeks. The last survey took place 12 weeks later (i.e., 1 week ahead of the final exam; T2). Participation in the surveys was voluntary and anonymous, and the surveys contained exercises like the ones on the exam (see also the Study 1 procedure). During the semester, we offered the same learning activities as in Study 1.

5.2 Sample

At the beginning of the semester, $N = 285$ preservice teachers and psychology undergraduates took part in an online survey. Of those, $n = 171$ took part in the survey at T2 and were therefore included in the analyses. Participants were on average $M = 21.6$ years old ($SD = 3.97$), and 77.2 % were female (22.3 % male and 0.5 % did not indicate their gender). They reported that they had been studying on average for $M = 3.59$ semesters ($SD = 2.74$); 56 % were preservice teachers and indicated studying a wide variety of subjects, whereas 44 % were psychology undergraduates. Missing data analyses revealed that students who continued to participate in our study had a better HSGPA ($M = 2.23$, $SD = 0.64$) than students who did not participate at T2 ($M = 1.91$, $SD = 0.66$), $t(281) = 3.88$, $p < .01$, $d = 0.49$. We computed equivalence tests, setting the equivalence bounds at $\Delta = 0.40$ because with at least 113 subjects in each condition (included and drop out) an effect of $d = 0.33$ can be detected with a probability of 80 % and is therefore considered the smallest effect size of interest here (see Lakens, 2017; Lakens et al., 2018). Equivalence tests indicated that on all other study variables, there were no further differences between students who participated at T2 and those who did not (all $p < .05$). Considering that 234 students completed the course and took part in the final exam, about 73 % of all students attending the lecture class participated in our study.

5.3 Measures

The measures were the same as in Study 1. Only this time, the mock exam used to assess *performance* at the end of the semester consisted of 55 items instead of 39, and we assessed ability self-concept with six instead of three items (*SESSKO*; Schöne et al., 2002).

5.4 Results

Again, we used R (version 4.0.4, R Core Team), the TOSTER package (Lakens, 2017) and IBM SPSS Statistics (version 24) for all other statistical analyses. Descriptive statistics and correlations of all study variables from Study 2 are presented in Tables 5 and 6. As in Study 1, students began the semester with high motivation and intentions to use all learning activities. The average HSGPA found in the sample represents a *good* performance according to the German grading system. The average prior knowledge score indicates that students began the lecture with little knowledge (because the mean score again represented the probability of guessing).

5.4.1 Intentions and use of learning activities

The activities students intended to use the most (see RQ1) were to attend the lecture regularly (M = 4.72, SD = 0.72) and to take practice tests regularly (M = 4.68, SD = 0.56). Paired t tests with an adjusted $\alpha < .007$ (Bonferroni-corrected) indicated that regular attendance of lecture sessions was significantly preferred over all other activities, $t(111) = 4.28$ to 19.77, all ps < .007, except taking practice tests regularly, $t(169) = 0.89$, $p = .377$, and memorizing content for the exam, $t(170) = 1.16$, $p = .247$. However, students intended to use many activities a lot (almost all intentions were above the scale mean).

Table 5

Study 2: Means and standard deviations for learning activities

(N = 171)	Intention (T1) M (SD)	Use (T2) M (SD)	Difference use-intention M (SD)
Memorizing content for the exam	4.63 (.69)	4.45 (.93)	–0.20 (.96)
Attending lecture regularly	4.72 (.72)	4.04 (1.29)	–0.68 (1.20)
Taking practice tests regularly	4.68 (.56)	4.47 (1.00)	–0.20 (1.04)
Reviewing lecture notes regularly	4.40 (.81)	3.87 (1.28)	–0.53 (1.20)
Submitting essays regularly	3.84 (.97)	2.43 (1.43)	–1.41 (1.47)
Forming a study group	3.59 (1.16)	2.85 (1.33)	–0.75 (1.24)
Submitting items regularly	3.64 (1.07)	1.45 (.85)	–2.19 (1.24)
Reviewing literature regularly	2.77 (1.08)	1.54 (.98)	–1.23 (1.26)

Note. Intentions to use and use of learning activities could be indicated on a scale ranging from 1 to 5.

Table 6

Study 2: Means, standard deviations, reliabilities, and bivariate correlations of study variables

(N = 171)	*M*	*SD*	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) HSGPA	3.23	(0.64)	-							
(2) Prior knowledge	8.20	(1.75)	0.05	-						
(3) Intentions (T1)	4.04	(0.50)	0.21**	–0.07	(.68)					
(4) Values (T1)	3.93	(0.67)	0.17*	–0.08	0.49**	(.89)				
(5) Self-concept (T1)	3.26	(.60)	0.07	–0.01	0.27**	0.50**	(.84)			
(6) Values (T2)	3.58	(0.78)	0.07	–0.10	0.39**	0.62**	0.34**	(.92)		
(7) Self-concept (T2)	3.09	(0.68)	0.18*	0.06	0.24**	0.37**	0.49**	0.54**	(.88)	
(8) Learning activities (T2)	3.13	(0.59)	0.32**	0.03	0.50**	0.24**	0.03	0.24**	0.24**	(.60)
(9) Performance	46.92	(5.67)	0.39**	0.15*	0.13	0.09	0.16*	0.01	0.20**	0.31**

Note. Self-report measures all used a response scale from 1 to 5, except for high school grade point average (HSGPA; possible range: 1 to 4); grades were recoded: larger numbers represent better grades. Prior knowledge could range from 0 to 15; performance could range from 0 to 55. Cronbach's alpha, when applicable, is listed on the diagonal.
* p < .05. ** p < .01.

To address RQ2 about the actual use of learning activities during the semester, we inspected the means and standard deviations of each learning activity. The activity used the most—on a descriptive level—was practice tests ($M = 4.43$, $SD = 0.88$). However, we again interpret this finding cautiously (see results from Study 1). Apart from the practice tests, students reported to have memorized the content for the exam above all other activities, $t(170) = 3.04$ to 29.85, all ps $< .007$. Table 5 shows the actual use of all activities. Looking at the corresponding intentions to use each learning activity, students used each activity less than they had intended (see Table 5). At the end of the semester, students reported that overall, they implemented their intentions to a lesser extent than intended, $t(170) = 21.45$, $p < .001$, $d = 1.64$. However, the more students intended to use the learning activities, the more they implemented the activities ($r = .50$, $p < .01$).

5.4.2 Predicting performance

Bivariate correlations (see Table 6) indicated that performance at T2 was related to HSGPA, ability self-concept at T2, and use of learning activities. In this sample, prior knowledge and ability self-concept at T1 were also related to performance.

We computed a multiple regression to explain performance at T2, entering all significant correlates except HSGPA in a first step, and adding HSGPA in a second step to test H1 and H2. In this sample, more of the assumed predictors correlated with performance than in Study 1, so the set of predictors differed slightly. Results are presented in Table 7. The learning activities again explained performance, even after controlling for HSGPA, thereby supporting H1 and H2 again. In the final model, all predictors together accounted for 20 % of the variance in performance.

Table 7

Study 2: Hierarchical regression predicting performance on the mock exam (z-transformed)

Variable	*B*	95% CI for *B*		*SE B*	ß	R^2	ΔR^2
		LL	*UL*				
Step 1						0.11	0.11**
Constant	–2.92**	–4.16	–1.67	0.63			
Prior knowledge	0.06	–0.02	0.14	0.04	0.11		
Self-concept (T1)	0.19	–0.09	0.46	0.14	0.11		
Self-concept (T2)	0.10	–0.14	0.36	0.13	0.07		
Learning activities (T2)	0.47**	0.23	0.72	0.13	0.28**		
Step 2						0.20	0.09**
Constant	–3.83**	–5.08	–2.58	0.64			
Prior knowledge	0.06	–0.02	0.13	0.04	0.10		
Self-concept (T1)	0.18	–0.08	0.44	0.13	0.11		
Self-concept (T2)	0.06	–0.18	0.29	0.12	0.04		
Learning activities (T2)	0.32*	0.07	0.56	0.13	0.19*		
HSGPA	0.49**	0.27	0.72	0.11	0.32**		

Note. N = 171. CI = confidence interval, HSGPA high school grade point average.

* $p < .05$. ** $p < .01$.

Goodman and Kruskal's Gamma correlations to investigate bivariate associations between each of the eight learning activities and performance on the mock exam were computed. Results are presented in Table 8.

Table 8

Study 2: Bivariate associations of single learning activities and performance on the mock exam.

(N = 171)	Memorizing content for the exam	Attending lectures regularly	Taking practice tests regularly	Reviewing lectures notes regularly	Submitting essays regularly	Forming a study group	Submitting items regularly	Reviewing literature regularly
Performance (T2)	0.22*	0.32**	0.13	0.23**	0.27**	0.09	0.00	–0.06

Note. Use of learning activities could be indicated on a scale ranging from 1 (*not at all*) to 5 (*very much*). Performance could range from 0 to 55. * $p < .05$, ** $p < .01$.

5.4.3 Predicting the use of learning activities

In addition to the hypotheses for replicating the findings from Study 1, in Study 2, we had further hypothesized that motivation at the beginning of the semester would predict the use of learning activities during the semester. In a multiple regression, motivational values at the beginning of the semester predicted the use of learning activities during the semester ($\beta = 0.29$, $p < .01$), whereas expectancies did not ($\beta = -0.11$, *ns*). The coefficient for motivational values did not change when controlling for prior achievement ($\beta = 0.29$, $p < .01$). Motivational values and prior achievement together explained 13 % of variance in the use of learning activities.

6. Discussion

Besides individual learning prerequisites and learning experiences that have already contributed to achievement in the past, students' engagement in self-regulated learning is essential for success in higher education (see, e.g., Zimmerman, 2008). We considered it useful to assess a proximal predictor of achievement and therefore focused on the behavioral aspect of self-regulated learning, namely, learning activities. To investigate the use and usefulness of specific learning activities, we designed a lecture class that included specific activities that relied on evidence-based principles such as spaced learning, self-testing (Dunlosky et al., 2013), and feedback (Downs, 2015).

6.1 Main findings

We assessed students' learning prerequisites, prior achievement, and use of these activities over the course of one semester as well as their performance at the end of the semester. Study 1 demonstrated that students' use of learning activities was associated with their success beyond their individual learning prerequisites (prior knowledge and motivation) and prior achievement, whereas prior achievement still predicted performance. In Study 2, we cross-validated this finding and further investigated the antecedents of students' use of learning activities. The more students valued the subject, the more they engaged in learning activities. However, expectancies were not associated with the use of learning activities. In both studies, we found a large gap between students' intentions and behavior.

Regarding the explanation of success in higher education, the studies add to existing research as follows: Prior achievement explained achievement in our studies, as has frequently been shown in the past (Kobrin et al., 2008; Richardson et al., 2012). This emphasizes the importance of considering prior achievement when investigating the incremental validity of new predictors of success in higher education. Nonetheless, our results support the idea that students' active engagement matters, and that they can actively make a difference in their own success. It has to be acknowledged that we only found small effects of the learning activities. However, these activities may help students compensate for heterogeneous learning prerequisites because the learning activities can be used by everyone. We therefore think of the learning activities as a valuable supplement to teaching in higher education. Furthermore, the activities described here, or similar activities that rely on the same principles, can be implemented in manifold learning situations in higher education and help instructors monitor students' knowledge level and adapt their teaching to it. Thereby the results provide instructors with ideas about how to optimize their teaching. As learning can be understood as an interaction of supply and use, this also points out how challenging it is for instructors to equip students with effective extracurricular activities.

Of course, there may be moderator variables that help to understand the relationship between students' activity use and performance in more detail. For further research, it may be promising to have a closer look at possible moderators of this relationship. From motivational research we know that quality of motivational regulation moderates the relationship between motivational regulation and performance (Engelschalk et al., 2017). For our learning activities, quality of activity use may also moderate the association of learning activity use with performance. For example, if students use study groups to quiz each other and give one another feedback, this study group probably benefits learning more than a study group in which students only spend time together rereading material. Further, students could benefit from different learning activities differently, for example according to their preferences. Another possible moderator variable in our design is students' conscientiousness and ambition. Students that are more conscientious or ambitious may invest more effort and time in each single learning activity and use them more effectively. Further research could benefit from including these variables and thereby help understanding the complex processes of learning. This may also help to increase the explained variance in performance which was rather small in our studies.

Although the main results of Study 1 could be replicated in Study 2, there were differences in the pattern of correlations. For example, prior knowledge was not associated with performance in Study 1 but it was in Study 2. Such differences in the observed correlations may be due to differences in the content of the lecture classes. One took place in the summer term and one in the winter term. Differences in the pattern of correlations maybe represent that there are always some inaccuracies in our data - especially when conducting research in the field. However, the main finding that prior achievement and learning activity use explained performance was found in both studies, regardless of the sample.

Further, we found not all activities to be associated with performance when considered individually. This may indicate that several learning activities are more helpful to students (e.g., submitting essays) than others (e.g., using study groups). However, this finding needs to be interpreted with caution. In order to compare the usefulness of activities, they should be assessed with more items.

The finding that a motivational construct explained learning behavior is in line with previous motivational research that has shown the expectancies and values that are associated with achievement-related behavior such as effort (Putwain et al., 2019). However, expectancies at the beginning of the semester were not related to the use of learning activities during the semester in our studies. From one point of view, this might sound surprising because the expectancy component of motivation has been shown to be more predictive when it comes to explaining achievement (Schneider & Preckel, 2017). However, it might be the case that expectancies affect achievement via mechanisms other than active engagement. It might even be the case that students with high expectancies do not consider it necessary to engage in optional activities. At the same time, the results of Study 2 reveal the importance of the value component of motivation. The more interesting, useful, and important students perceive a subject to be, the more they engage. Instructors could address this in lectures by highlighting the usefulness and importance of a subject and by encouraging students to search for personally interesting, useful, and important aspects of the subject (see, e.g., Yeager et al., 2014).

However, high values seem insufficient to ensure engagement. The large gap between intentions and behavior has been found in many domains of self-regulated behavior such as health behavior (e.g., Sniehotta et al., 2005). In this context, however, this shows how important it is for students to maintain their effort over the course of a longer period of time (e.g., 1 semester). Because there is inter-individual variation in the difference between intentions and use of learning activities in our sample, further research could use difference-scores as another measure of (problems with) self-regulation or students' ability to realistically set their intentions and relate this to academic outcomes. Of course, researchers could also figure out factors that help students put their intentions into practice. One possibility is to look at factors from the theory of planned behavior (Ajzen, 1991). Perhaps some students did not feel in control of engaging in the activities or did not feel a subjective norm to use the activities. These factors could be considered in future research. Another possibility could be to make learning activities more attractive, for example, by showing students the results of studies—such as ours—that underline their usefulness or design interventions that help students to regulate their learning. Some promising interventions have already been investigated, including web-based trainings for self-regulated learning (Bellhäuser et al., 2016) and interventions including implementation intentions regarding the use of optional activities (van der Beek et al., 2020).

6.2 Limitations

Of course, our study has some limitations. First, we had to rely on self-report data from our students. The assessment of learning activities could be improved by assessing them automatically (e.g., essay submissions) or more frequently. In this study, we relied on self-report data due to the importance of anonymous data collection. Nevertheless, we think students' report of their learning activities should have been quite valid because participation in our surveys was voluntary. Furthermore, we asked about specific behavior, and this should have helped the students give correct estimates. In the future, it would be worthwhile to work on a more objective but still anonymous way to assess these activities. Second, we only investigated these activities in one lecture class and cannot generalize to other learning situations. However, we did cross-validate our finding in another sample of students. Furthermore, the learning activities we described relied on general principles (e.g., practice testing) and could be adapted to other contexts. In our studies, we had participation rates of 62 % (Study 1) and 73 % (Study 2) of all students who took the final exam. Therefore, our results could be biased because more motivated students participated in more surveys. However, we still had some variability in our sample regarding motivation to explain achievement and learning behavior. Furthermore, dropout analyses indicated that students who participated in both surveys did not differ in motivation from those who dropped out. Of course, we cannot make causal inferences because the study lacks an experimental design. However, by implementing a longitudinal design and controlling for other strong and well-established predictors of success, we have good reasons to assume that students benefit from the offered learning activities. Further moderator variables as described above may help to further clarify the associations between learning activity use and achievement. The consideration of further explaining variables may also improve the variance explained in performance. In our study, we could only explain up to 20 % of variance in performance and regression coefficients were small to moderate. However, as we found these associations in a realistic field setting, they are nonetheless noteworthy.

Another limitation is the sample size we could investigate. With more students participating in the online surveys the analyses' power would increase and maybe also help to give more accurate estimates of the size of correlations of regression coefficients. The correlation coefficients differ in their size. However, the main associations are comparable between samples. In the field, the sample size oftentimes is limited. The learning activities were implemented in one lecture class and therefore the sample is limited to the students attending this lecture. Nonetheless, future studies may benefit from implementing these activities into more and different lecture classes, thereby increasing the sample size and at the same time investigating the generalizability of the results.

6.3 Implications

With these limitations in mind, the present studies revealed interesting and important results for theory and practice. The results support the idea that students' self-regulation matters for students' success in higher education. We specified self-regulated learning by investigating specific learning activities in the field and over the course of a semester and thereby improved the assessment of self-regulated learning. At the same time, we described and evaluated useful ideas for teaching practice in higher education. Learning in higher education is an interaction of supply and use: Instructors who supply students with effective learning activities and students who use these activities. Of course, students' learning prerequisites and prior achievement still contribute to explaining their success. However, beyond this, students can actively make a difference in their success by using learning activities. Looking at the gap between intentions and learning activity use, further research could investigate this gaps' relevance for students' success and consider interventions that help students to improve their learning even more.

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179-211.

Bae, C. L., Therriault, D. J., & Redifer, J. L. (2019). Investigating the testing effect: Retrieval as a characteristic of effective study strategies. *Learning and Instruction*, *60*, 206–214. https://doi.org/10.1016/j.learninstruc.2017.12.008

Beek, S. V. D., Bellhäser, H., & Hertel, S. (2020). Do Minimal Interventions Increase Participation Rates in Voluntary Online Training at High School?. *Psychology Learning & Teaching, 0*(0), 1-23. https://doi.org/10.1177/1475725720965002

Bellhäuser, H., Lösch, T., Winter, C., & Schmitz, B. (2016). Applying a web-based training to foster self-regulated learning—Effects of an intervention for large numbers of participants. *The Internet and Higher Education*, *31*, 87-100. https://doi.org/10.1016/j.iheduc.2016.07.002

Blasiman, R. N., Dunlosky, J., & Rawson, K. A. (2017). The what, how much, and when of study strategies: comparing intended versus actual study behaviour. *Memory*, *25*(6), 784-792. https://doi.org/10.1080/09658211.2016.1221974

Boekaerts, M., Pintrich, P., & Zeidner, M. (2000). Handbook of self-regulation. San Diego, CA: Academic Press.

Boser, J., Scherer, S., Kuchta, K., Wenzel, S. F. C., & Horz, H. (2017). Empirically founded teaching in psychology–An example for the combination of evidence-based teaching and the scholarship of teaching and learning. *Psychology Learning & Teaching*, *16*(2), 261-275. https://doi.org/10.1177/1475725716686452

Brandstätter, H., & Mayr, J. (1994). Die "Lehrer-Persönlichkeits-Adjektivskalen" (LPA). Ein Instrument zur Selbsteinschätzung berufsrelevanter Persönlichkeitsmerkmale. [Teacher Personality Adjectives Scales]. In J. Mayr (Ed.), *Lehrer/in werden* (pp. 231-247). Innsbruck: Österreichischer Studienverlag.

Breitwieser, J., Neubauer, A., Schmiedek, F., & Brod, G. (2020). Self-Regulation prompts promote the achievement of learning goals–but only briefly: Uncovering hidden dynamics in the effects of a psychological intervention. PsyArxiv. https://doi. org/10.31234/osf.io/bz49m.

Brühl, B., Döpfner, M., & Lehmkuhl, G. (2000). Der Fremdbeurteilungsbogen für hyperkinetische Störungen (FBB-HKS) - Prävalenz hyperkinetischer Störungen im Elternurteil und psychometrische Kriterien [parent and teacher rating scale assessing hyperkinesis - Prevalence of hyperkinesis according to parent ratings and psychometric properties]. *Kindheit und Entwicklung, 9*(2), 115-125.

Brühwiler, C., & Blatchford, P. (2011). Effects of class size and adaptive teaching competency on classroom processes and academic outcome. *Learning and Instruction*, *21*(1), 95-108. https://doi.org/10.1016/j.learninstruc.2009.11.004

Carpenter, S. K. (2012). Testing enhances the transfer of learning. *Current Directions in Psychological Science*, *21*(5), 279–283. https://doi.org/10.1177/0963721412452728

Carpenter, S. K., Pashler, H., Wixted, J. T., & Vul, E. (2008). The effects of tests on learning and forgetting. *Memory & Cognition*, *36*(2), 438-448. https://doi.org/10.3758/MC.36.2.438

Cogliano, M., Kardash, C. M., & Bernacki, M. L. (2019). The effects of retrieval practice and prior topic knowledge on test performance and confidence judgments. *Contemporary Educational Psychology*, *56*, 117–129. https://doi.org/10.1016/j.cedpsych.2018.12.001

Coltman, T., Devinney, T. M., Midgley, D. F., & Venaik, S. (2008). Formative versus reflective measurement models: Two applications of formative measurement. *Journal of Business Research*, *61*(12), 1250-1262. https://doi.org/10.1016/j.jbusres.2008.01.013

Credé, M., & Kuncel, N. R. (2008). Study habits, skills, and attitudes: The third pillar supporting collegiate academic performance. *Perspectives on Psychological Science*, *3*(6), 425–453. https://doi.org/10.1111/j.1745-6924.2008.00089.x

Credé, M., & Phillips, L. A. (2011). A meta-analytic review of the Motivated Strategies for Learning Questionnaire. *Learning and Individual Differences*, *21*(4), 337–346. https://doi.org/10.1016/j.lindif.2011.03.002

Credé, M., Roch, S. G., & Kieszczynka, U. M. (2010). Class attendance in college: A meta-analytic review of the relationship of class attendance with grades and student characteristics. *Review of Educational Research*, *8*(2), 272–295. https://doi.org/10.3102/0034654310362998

Dochy, F., Segers, M., & Buehl, M. M. (1999). The relation between assessment practices and outcomes of studies: The case of research on prior knowledge. *Review of Educational Research*, *69*(2), 145-186. https://doi.org/10.3102/00346543069002145

Downs, S. D. (2015). Testing in the college classroom: Do testing and feedback influence grades throughout an entire semester? *Scholarship of Teaching and Learning in Psychology, 1*(2), 172–181. https://doi.org/10.1037/stl0000025

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, *14*(1), 4–58. https://doi.org/10.1177/1529100612453266

Dunn, D. S., Saville, B. K., Baker, S. C., & Marek, P. (2013). Evidence-based teaching: Tools and techniques that promote learning in the psychology classroom. *Australian Journal of Psychology*, *65*(1), 5-13. https://doi.org/10.1111/ajpy.12004

Dutke, S., & Barenberg, J. (2015). Easy and informative: Using confidence-weighted true–false items for knowledge tests in psychology courses. *Psychology Learning & Teaching*, *14*(3), 250-259. https://doi.org/10.1177/1475725715605627

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives* (pp. 75–146). San Francisco: Freeman.

Engelschalk, T., Steuer, G., & Dresel, M. (2017). Quantity and quality of motivational regulation among university students. *Educational Psychology*, *37*(9), 1154-1170.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175-191. https://doi.org/10.3758/BF03193146

Geller, J., Toftness, A. R., Armstrong, P. I., Carpenter, S. K., Manz, C. L., Coffman, C. R., & Lamm, M. H. (2018). Study strategies and beliefs about learning as a function of academic achievement and achievement goals. *Memory*, *26*(5), 683-690. https://doi.org/10.1080/09658211.2017.1397175

Gollwitzer, P. M. (1996). The volitional benefits of planning. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action: Linking cognition and motivation to behavior* (pp. 287-312). New York: Guilford

Gottfried, A. E., Fleming, J. S., & Gottfried, A. W. (2001). Continuity of academic intrinsic motivation from childhood through late adolescence: A longitudinal study. *Journal of Educational Psychology, 93*(1)*,* 3-13. https://doi.org/10.1037//0022-0663.93.1.3

Hartwig, M. K., & Dunlosky, J. (2012). Study strategies of college students: Are self-testing and scheduling related to achievement? *Psychonomic Bulletin & Review*, *19*(1), 126–134. https://doi.org/10.3758/s13423-011-0181-y

Hattie, J. (2011). Which strategies best enhance teaching and learning in higher education? In D. Mashek, & E.Y. Hammer (Eds.), *Empirical research in teaching and learning: Contributions from Social Psychology* (pp.130–142). London: Blackwell Publishing.

Helmke, A. (2017). Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts [Teaching quality and teacher professionality. Diagnosis, evaluation, and improvement of teaching] (7th ed.). Friedrich Verlag.

Jansen, R. S., Van Leeuwen, A., Janssen, J., Jak, S., & Kester, L. (2019). Self-regulated learning partially mediates the effect of self-regulated learning interventions on achievement in higher education: A meta-analysis. *Educational Research Review*, *28*, 100292. https://doi.org/10.1016/j.edurev.2019.100292

Jonsson, B., Wiklund-Hörnqvist, C., Stenlund, T., Andersson, M., & Nyberg, L. (2020). A learning method for all: The testing effect is independent of cognitive ability. *Journal of Educational Psychology.* Advance online publication. https://doi.org/10.1037/edu0000627

Karpicke, J. D., Butler, A. C., & Roediger III, H. L. (2009). Metacognitive strategies in student learning: Do students practise retrieval when they study on their own? *Memory*, *17*(4), 471-479. https://doi.org/10.1080/09658210802647009

Kim, Y. E., Brady, A. C., & Wolters, C. A. (2020). College students' regulation of cognition, motivation, behavior, and context: Distinct or overlapping processes? *Learning and Individual Differences*, *80,* 101872. https://doi.org/10.1016/j.lindif.2020.101872

Kobrin, J. L., Patterson, B. F., Shaw, E. J., Mattern, K. D., & Barbuti, S. M. (2008). *Validity of the SAT for predicting first-year college grade point average.* New York, NY: College Board.

Kornell, N., & Bjork, R. A. (2007). The promise and perils of self-regulated study. *Psychonomic Bulletin & Review*, *14*(2), 219–224. https://doi.org/10.3758/BF03194055

Köller, O. (2012). What works best in school? Hatties Befunde zu Effekten von Schul- und Unterrichtsvariablen auf Schulleistungen [What works best in school? Hattie's findings on effects of school and teaching variables on academic achievement]. *Psychologie in Erziehung und Unterricht*, *59*, 72-78. http://dx.doi.org/10.2378/peu2012.art06d

Kuncel, N. R., & Hezlett, S. A. (2007). Standardized tests predict graduate students' success. *Science, 315*(5815), 1080–1081. https://doi.org/10.1126/science.1136618

Lakens, D. (2017). Equivalence tests: A practical primer for t tests, correlations, and meta-analyses. *Social psychological and personality science*, *8*(4), 355-362. https://doi.org/10.1177/1948550617697177

Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence Testing for Psychological Research: A Tutorial. *Advances in Methods and Practices in Psychological Science*, *1*(2), 259-269. https://doi.org/10.1177/2515245918770963

Larose, S.A., Bernier, A., & Tarabulsy, G.M. (2005). Attachment state of mind, learning dispositions, and academic performance during the college transition. *Developmental Psychology, 41*(1), 281–289. https://doi.org/10.1037/0012-1649.41.1.281

LeBreton, J. M., Hargis, M. B., Griepentrog, B., Oswald, F. L., & Ployhart, R. E. (2007). A multidimensional approach for evaluating variables in organizational research and practice. *Personnel Psychology*, *60*(2), 475–498. https://doi.org/10.1111/j.1744-6570.2007.00080.x

Liborius, P., Bellhäuser, H., & Schmitz, B. (2019). What makes a good study day? An intraindividual study on university students' time investment by means of time-series analyses. *Learning and Instruction*, *60*, 310-321. https://doi.org/10.1016/j.learninstruc.2017.10.006

Mayr, J. (1998). Die „Lehrer-Interessen-Skalen“ (LIS). Ein Instrument für Forschung und Laufbahnberatung. [Teacher Interest Scales: An Instrument for Research and Vocational Guidance]. In J. Abel & C. Tarnai (Eds.), *Pädagogisch-psychologische Interessenforschung in Studium und Beruf* (pp. 111-125). Münster: Waxmann.

Molenaar, I., Horvers, A., & Baker, R. S. (2019). What can moment-by-moment learning curves tell about students' self-regulated learning?. *Learning and Instruction*, 101206. https://doi.org/10.1016/j.learninstruc.2019.05.003

Morisano, D., Hirsh, J. B., Peterson, J. B., Pihl, R. O., & Shore, B. M. (2010). Setting, elaborating, and reflecting on personal goals improves academic performance. *Journal of Applied Psychology*, *95*(2), 255.

Musu-Gillette, L. E., Wigfield, A., Harring, J. R., & Eccles, J. S. (2015). Trajectories of change in students' self-concepts of ability and values in math and college major choice. *Educational Research and Evaluation*, *21*(4), 343–370. https://doi.org/10.1080/13803611.2015.1057161

OECD (2018), *Education at a Glance 2018: OECD Indicators*, OECD Publishing, Paris, https://doi.org/10.1787/eag-2018-en.

OECD (2019), *Education at a Glance 2019: OECD Indicators*, OECD Publishing, Paris, https://doi.org/10.1787/f8d7880d-en.

Pashler, H., Bain, P. T., Bottge, B., Koedinger, K., McDaniel, M., & Metcalf, J. (2007). *Organizing instruction and study to improve student learning*. Washington, DC: National Center for Education Research, Institute of Education Science, U.S., Department of Education.

Perry, R. P., Hladkyj, S., Pekrun, R. H., & Pelletier, S. T. (2001). Academic control and action control in the achievement of college students: A longitudinal field study. *Journal of Educational Psychology*, *93*(4), 776.

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review*, *16*(4), 385-407. https://doi.org/1040-726X/04/1200-0385/0

Pintrich, P. R., Smith, D. A., Garcia, 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. https://doi.org/10.1177/0013164493053003024

Putwain, D. W., Nicholson, L. J., Pekrun, R., Becker, S., & Symes, W. (2019). Expectancy of success, attainment value, engagement, and achievement: A moderated mediation analysis. *Learning and Instruction, 60,* 117–125. https://doi.org/10.1016/j.learninstruc.2018.11.005

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://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. https://doi.org/10.1037/a0026838

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin*, *130*(2), 261-288. https://doi.org/10.1037/0033-2909.130.2.261

Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, *15*(1), 20-27. https://doi.org/10.1016/j.tics.2010.09.003

Rowland, C. A. (2014). The effect of testing versus restudy on retention: A meta-analytic review of the testing effect. *Psychological Bulletin*, *140*(6), 1432-1463. http://dx.doi.org/10.1037/a0037559

Schaller, P. & Spinath, B. (2017). Selbstberichtskalen zur Erfassung motivationsbezogener Kompetenzen (MOBEKO) im Studium: Entwicklung und Validierung [A self-report instrument for measuring motivation-based competences (MOBEKO) of students in higher education: Development and validation.]. *Diagnostica*, *63*(3), 229-241. http://dx.doi.org/10.1026/0012-1924/a000176

Schneider, M., & Preckel, F. (2017). Variables associated with achievement in higher education: A systematic review of meta-analyses. *Psychological Bulletin*, *143*(6), 565-600. http://dx.doi.org/10.1037/bul0000098

Schöne, C., Dickhäuser, O., Spinath, B., & Stiensmeier-Pelster, J. (2002). *Skalen zur Erfassung des schulischen Selbstkonzepts: SESSKO* [Scales for measuring academic ability self-concept]. Hogrefe.

Seidel, T. (2014). Angebots-Nutzungs-Modelle in der Unterrichtspsychologie. Integration von Struktur-und Prozessparadigma. *Zeitschrift für Pädagogik*, *60*(6), 850-866. URN: urn:nbn:de:0111-pedocs-146864

Seifried, E., Eckert, C., & Spinath, B. (2018). Optional learning opportunities: Who seizes them and what are the learning outcomes?. *Teaching of Psychology*, *45*(3), 246-250. https://doi.org/10.1177/0098628318779266

Seifried, E., & Spinath, B. (2020). Using formative self-assessment to improve teaching and learning in educational psychology courses. In E. Landrum, C. Hakala, & S. Nolan (Eds.), *Assessing Undergraduate Learning in Psychology: Strategies for Measuring and Improving Student Performance* APA Books.

Sniehotta, F. F., Scholz, U., & Schwarzer, R. (2005). Bridging the intention–behaviour gap: Planning, self-efficacy, and action control in the adoption and maintenance of physical exercise. *Psychology & Health*, *20*(2), 143–160. https://doi.org/10.1080/08870440512331317670

Spinath, B., Seifried, E., & Eckert, C. (2014). Forschendes Lehren: Ein Ansatz zur kontinuierlichen Verbesserung von Hochschullehre [Investigating teaching: An approach to the continuous improvement teaching in higher education]. *Journal Hochschuldidaktik, 25*, 14-16.

Steinmayr, R., & Spinath, B. (2010). Konstruktion und erste Validierung einer Skala zur Erfassung subjektiver schulischer Werte (SESSW) [Construction and First Validation of a Scale for Assessing Subjective School-related Values]. *Diagnostica, 56*, 195–211. https://doi.org/10.1026/0012-1924/a000023

Thompson, R. A., & Zamboanga, B. L. (2004). Academic aptitude and prior knowledge as predictors of student achievement in introduction to psychology. *Journal of Educational Psychology*, *96*(4), 778-784. https://doi.org/10.1037/0022-0663.96.4.778

Weinstein, C., Palmer, A. C., & Schulte, D. R. (1987). *Learning and Study Strategy Inventory (LASSI*). Clearwater, FL: H&H Publishing.

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, *25*(1), 68–81. https://doi.org/10.1006/ceps.1999.1015

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Hillsdale: Lawrence Erlbaum.

Yeager, D. S., Henderson, M. D., Paunesku, D., Walton, G. M., D'Mello, S., Spitzer, B. J., & Duckworth, A. L. (2014). Boring but important: a self-transcendent purpose for learning fosters academic self-regulation. *Journal of Personality and Social Psychology*, *107*(4), 559. https://doi.org/10.1037/a0037637

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). New York: Academic Press.

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal*, *45*(1), 166–183. https://doi.org/10.3102/0002831207312909

Paper II

This is the accepted version of the article

Bosch, Eva, & Spinath, Birgit (2023). Students' motivation in an online and a face-to-face semester: A comparison of initial level, development and use of learning activities. *Zeitschrift für Psychologie, 231(2), 93-102. https://doi.org/10.1027/2151-2604/a000519*

CRediT author statement

Eva Bosch: conceptualization, investigation, data curation, formal analysis, writing – original draft.

Birgit Spinath: conceptualization, supervision, investigation, writing – review & editing.

Students' Motivation in an Online and a Face-To-Face Semester: A Comparison of Initial Level, Development and Use of Learning Activities

Eva Bosch[1*] & Birgit Spinath[1]

[*]corresponding author; Eva.Bosch@psychologie.uni-heidelberg.de

[1]Heidelberg University, Department of Psychology, Hauptstraße 47-51, 69117 Heidelberg, Germany

Abstract

Challenges for university students were high during distance education in lockdowns due to the COVID19 pandemic. Self-regulation and motivation became more important but motivation was possibly challenged more. To investigate motivational differences and possible positive effects of evidence-based learning activities we followed two cohorts of preservice teachers over the course of one semester. One cohort was followed in 2019 in a face-to-face semester ($N_{2019} = 225$) and the second cohort was followed one year later during the first lockdown ($N_{2020} = 311$). Students indicated their motivation at five measurement occasions and reported their use of learning activities twice. Multigroup linear change models indicated an overall decline of motivation in both cohorts. Surprisingly, neither initial motivation level nor motivational change differed between cohorts. Students who used more learning activities reported a more positive motivational development. This highlights the chance of evidence-based learning activities for students' motivation in regular and distance education.

Keywords: COVID19 pandemic, Higher Education, Motivation, Evidence-Based Learning Activities

1. Introduction

In March 2020, the lockdown due to the COVID-19 pandemic challenged learners and teachers to adapt to distance education. First empirical research indicates that distance education came along with deleterious effects for students' well-being (Kedraka & Kaltsidis, 2020; Schwinger et al., 2020; Steinmayr et al., 2022). In higher education, many instructors converted their courses into online formats, and, for example, made self-study materials available to students online. This required a high degree of self-regulation on the part of the students when they wanted to learn successfully. At the same time, students reported problems starting online classes on their own and staying focused (Wong, 2020). Correspondingly, Lockl et al. (2021) reported that distance learning required students to engage in higher levels of self-regulation.

In this context, motivation to learn is particularly important, because it is a crucial prerequisite for self-regulated learning (see e.g., Pintrich, 2004). To self-regulate learning successfully, students need motivation to initiate learning and also to maintain motivation over time. However, several researchers have found students' motivation to decline across different developmental phases and time spans (e.g., Benden & Lauermann, 2021; Jacobs et al., 2002; Kosovich et al., 2017; Robins et al., 2019). In online courses during pandemic-related lockdowns, motivation was possibly threatened more, because many prerequisites for motivation were missing (e.g., social presence and interaction, see also Müller et al., 2021) and students reported an increased amount of stress (Usher et al., 2021). The question arises as to what differences emerge in students' motivation during distance learning compared to face-to-face learning.

Extending the question of motivational challenges during distance education, it is important to examine what helps students stay motivated, be it in distance learning or face-to-face learning. One promising idea to help students stay engaged and motivated is offering evidence-based learning activities. Given the challenges of pandemic-related distance learning,

learning activities that can be easily implemented in an online course, in particular, offer an opportunity to facilitate distance learning for students. Activities that help students space out their learning, test their knowledge repeatedly and receive feedback on their performance have already been shown to boost students' performance (e.g., Dunlosky et al., 2013). Relying on a Supply-Use-Model, prior research has shown that, in a regular semester, the use of such activities explained students' performance even beyond prior achievement and motivation (Bosch et al., 2021). The present study investigates whether the use of such activities is associated with a more favorable motivational development, both in a regular and in a distance learning context.

2. Student Motivation

In our research, we are interested in university students' intrinsic motivation in terms of the intrinsic value component from Expectancy-Value Theory (EVT; Eccles et al., 1983; Wigfield & Eccles, 2000). Intrinsic motivation describes the pleasure and interest with which one approaches a topic or task. Together with other motivational values and performance expectations it is supposed to influence achievement related behavior and choices (Eccles & Wigfield, 2020). Intrinsic motivation is important in education for two reasons. First, empirical research shows that students who have high intrinsic motivation and performance expectancy (expectancy component of EVT) tend to learn more and perform better (e.g., Kriegbaum et al., 2018; Robinson et al., 2019; Trautwein et al., 2012). They also tend to drop out from university courses less frequently (Benden & Lauermann, 2021). Second, intrinsic motivation is also a desirable outcome of learning processes, as students should leave higher education courses with high motivation to learn and strong interest in the field.

Unfortunately, students suffer from motivational challenges. Longitudinal studies show that students' motivational values for various school subjects decrease overall from first to twelfth grade (e.g., Jacobs et al., 2002). A negative trend in student motivation was also found through the first two years of college in one study (Robins et al., 2019). Even over a relatively short period of time such as an academic semester, student motivation develops negatively (Benden & Lauermann, 2021; Kosovich et al., 2017). Negative trends in motivation can have negative consequences for students. One serious consequence of too little motivation is when students fail to achieve their goals and drop courses (Benden & Lauermann, 2021) or even university. In Germany about 25 % of students drop out from university before completing their bachelor's degree (Heublein & Wolter, 2011; OECD, 2013) and lack of motivation was identified as an important predictor of dropout, along with poor performance and financial problems (Heublein & Wolter, 2011).

In summary, there are many reasons why students should stay motivated. However, they seem to face motivational challenges during their academic careers. During the lockdown and distance learning, more challenges probably came along. Demands on students have increased, as they should take more responsibility for their learning process and become more self-directed learners (Lockl et al, 2021). They had to independently initiate learning sessions, stay focused alone in front of their electronic device, and structure their day. Not all students were able to cope well with these requirements (Lockl et al., 2021; Wong, 2020). With increased challenges and stress, students' motivation could have suffered more during distance education. Under an EVT perspective, Wang and Eccles (2012) investigated classroom characteristics that facilitate students' intrinsic value for a task. They found those characteristics helpful for students' intrinsic task value that support feelings of competence, connectedness, and autonomy, concepts from self-determination theory (SDT; Ryan & Deci, 2020). According to SDT, a crucial prerequisite for motivation is satisfaction of the basic psychological needs for relatedness, autonomy, and competence (Ryan & Deci, 2020). Because social interaction was diminished during distance education, the need for relatedness was possibly frustrated in many students. Wong (2020), for example, found that especially learners' need for relatedness was frustrated in online learning but not autonomy and competence. Such a frustration of one need may contribute to a decreased motivation in distance education. However, because temporal and spatial flexibility were increased in online courses, the need for autonomy could have been satisfied more at the same time. How such frustration of one need (relatedness) can be compensated by increased satisfaction of another need (autonomy), and how the different needs differ in strength and importance at the individual level are as yet open questions (e.g., Vansteenkiste et al., 2020). However, given the many changes and challenges students faced and the increasing stress, it would stand to reason that their motivation suffered. Indeed, first empirical research suggests detrimental effects of pandemic-related distance education on university student motivation. In an exploratory study Usher and colleagues (2021) retrospectively surveyed *N* = 358 U.S. university students at the end of the spring 2020 semester about their experiences and motivational changes. The majority of students reported that they procrastinated more and were less motivated during lockdowns. In another study, students retrospectively reported that courses that were moved online due to the pandemic were less enjoyable and less interesting (Garris & Fleck, 2022). However, the authors could not compare motivation before and after the change to distance education directly. Müller and colleagues (2021) surveyed two cohorts of students, one before and one during forced distance learning. Using a self-determination approach, they found that less desirable forms of motivation increased during lockdown. However, these studies could not look at the development of students' motivation over time. We are not aware of any study to date that has directly compared the intrinsic motivation and its' development of higher education students before and after switching to distance education.

The well-documented decline in motivation over time, as well as additional challenges during lockdowns and the potentially increased threat to motivation as a result, lead to the question of what might be done to mitigate negative motivation trends.

2.1 Antecedents of Motivation

What could help students staying motivated during regular and online semesters? Motivation is most often looked at as an antecedent of behavior, but behavior is also an antecedent of motivation, as described in several theoretical frameworks. In the model of situated EVT from Eccles and Wigfield (2020) for example, previous achievement-related experiences are thought to influence motivational values. Also, in models of self-regulated learning (e.g., Schmitz & Wiese, 2006) every learning action is followed by a reflection of outcomes. The results of this reflection are then integrated to build motivation for the next learning action. Hidi and Renninger (2006) argue that evoking situational interest makes students engage repeatedly, thereby strengthening individual interest in a certain domain more generally.

In their meta-analytic review, Credé and Kuncel (2008) investigated predictors of academic success in higher education and identified study habits, skills, and attitudes as an important pillar for academic success. Both, motivation as well as study skills and habits explained performance beyond cognitive predictors. But how these two predictors interact, is not that clear, yet. Interestingly, Credé and Kuncel (2008) propose a model of academic success that implies behavior (study habits and attitudes) as an antecedent of motivation (p. 430). This relationship, however, is not further specified or investigated empirically. In another very recent review of meta-analyses, Jansen et al. (2022) investigated the antecedents of K-12 students' motivation. They report that students' previous achievement, as well as instructional practices, were related to students' motivation with medium order effect sizes. However, they also note that the antecedents of higher education students' motivation have not been as widely studied. As mentioned earlier, a key challenge for higher education students is *maintaining* motivation over time in a self-directed learning environment, and this challenge can be even more demanding in distance education.

Taking behavior as an antecedent of motivation in a self-regulated learning context, we hypothesize that students' active learning behavior is associated with sustained motivation over time. In particular, we expect regular participation in learning activities to help students maintain intrinsic motivation in the subject.

2.2 Evidence-based Learning Activities

What could instructors do to support students' engagement and motivation? According to the Supply-Use Model for learning and instruction in higher education (see figure 1), instructors should provide learning opportunities that students can use to acquire knowledge (Bosch et al., 2021). When teachers think about which learning opportunities to offer, they should rely on evidence-based activities that help students to learn effectively (Dunn et al., 2013). The typical design of a lecture class includes regular lecture sessions. If students regularly attend sessions, they tend to receive better grades in university (Credé et al., 2010). Next to lecture sessions, empirical research has found that learning activities that help students space out their learning, test their knowledge repeatedly, and receive feedback on their performance promote student learning and performance (Downs, 2015; Dunlosky et al., 2013). For example, practice tests of the learning content allow students not only to assess their learning but enhance students' memory for the learned and for new material (e.g., Carpenter, 2012). The effect of testing is even stronger, when students receive feedback on their performance (Phelps, 2012). Much research on testing focuses on multiple-choice knowledge tests (e.g., Cogliano et al., 2019). Other useful test activities include writing assignments that require students to apply, integrate, and reflect on the content (Balgopal et al., 2012; Dunn et al., 2013). Further, the principle of distributed practice is related to students' performance (Dunlosky et al., 2013). Consequentially, activities should be effective if they allow students to space their learning over a period of time.

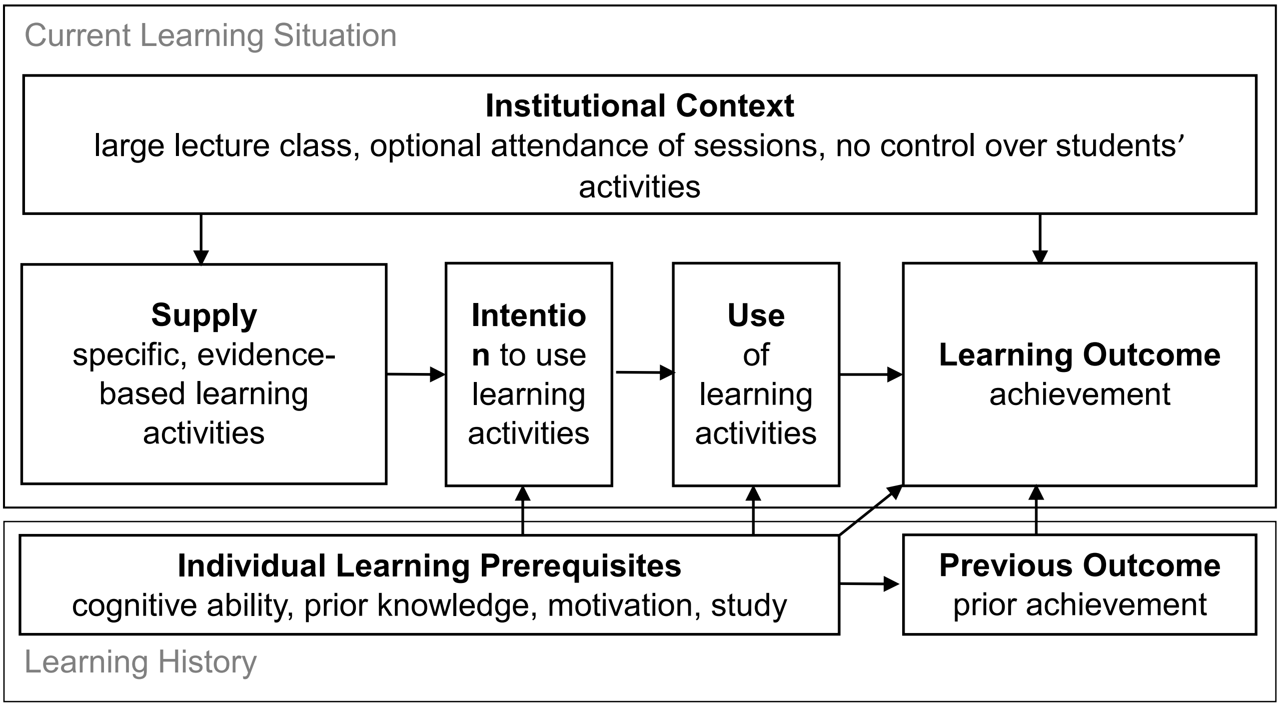


Figure 1. A Supply-Use Model of Learning, adapted for higher education (Bosch et al., 2021)

Based on these empirical results, a university lecture class was designed to promote student engagement. The lecture class included regular practice tests distributed over time in the form of knowledge-tests with confidence-weighted true-false items (see Dutke & Barenberg, 2015), writing assignments and feedback on these test activities. Such a lecture design has already been investigated empirically in the field and students who engaged in more activities performed better at the end of the course (Bosch et al., 2021). However, regular engagement in these evidence-based activities may not only improve students' knowledge but also help them to stay motivated by helping them discover interesting and joyful aspects of the content.

3. Summary and Research Questions

First empirical evidence suggests that students' learning motivation has suffered during distance education in the COVID-19 pandemic. The aim of the present research is to compare students' motivational level at the beginning of a semester and the development over the course of one semester in a regular face-to-face and an online lecture class. Given the motivational challenges in self-regulated learning contexts, we further investigate evidence-based learning activities as a tool that may help students stay engaged and motivated even in an online lecture class.

- **H1**: We expect to replicate the finding that students' intrinsic motivation declines over the course of one semester regardless of the lecture design (present or online).
- **H2**: We further expect that initial intrinsic motivation is lower and the decline is steeper in an online semester than in a regular face-to-face semester.
- **H3**: We hypothesize that students who more often attend lecture sessions and use evidence-based learning activities (practice tests in the form of self-tests and writing assignments) have smaller motivational declines.

Exploratively, we investigate which learning activities may be particularly important and offer a chance to improve teaching and learning, especially in online courses. Moreover, we analyzed which was the better predictor of students' dropout: use of learning activities or motivation early in the semester.

4. Method

To test the hypotheses stated above, self-report data from two cohorts of preservice teachers were analyzed and compared.

4.1 Procedure

The setting of the study was an introductory lecture class on educational psychology for preservice teachers at a German university. This lecture class is obligatory for preservice teachers. Cohort 1 was followed during summer term 2019 in a regular face-to-face semester. When the pandemic-related lockdown in March 2020 forced teachers and learners to convert to distance education, the design of the lecture class was left the same except for the weekly lecture sessions which switched to online. Cohort 2 was followed during this online term in summer 2020.

In both cohorts, lecture content and evidence-based learning activities were the same. The lecture class included weekly lecture sessions (face to face 2019 and online 2020) as well as practice tests in the form of self-tests and writing assignments online. All activities were optional. However, we recommended students to engage in as many activities as possible. On five self-tests, students were quizzed on the lectures' topics of the last three weeks with confidence-weighted true-false items (Dutke & Barenberg, 2015) and received feedback on their performance (number of correctly answered questions). With up to eight writing assignments, students had the opportunity to reflect on the content of the lecture and apply it to a new problem. Students received individual feedback on these assignments. The practice tests were offered every week during the semester and thereby helped students learn continuously and space their learning. The design of the learning activities is described in further detail by Seifried and Spinath (2020).

For data collection in both cohorts, all students who registered for the course were invited to complete online surveys after each self-test, starting after the first lecture session and from this point on every three weeks. Each time students were asked to indicate their intrinsic motivation. In the middle (T3) and at the end of the semester (T5), class attendance and use of optional evidence-based learning activities were assessed.

At several time points, we further assessed scales that were related to the evaluation of the lecture or illustration of the lectures' content and thus were not used in the present study. Participation in the surveys was voluntary and anonymous (students generated anonymous codes for longitudinal monitoring). Students were informed about the content and aims of the survey and that they would have no disadvantage if they did not participate.

4.2 Sample

At the beginning of the semester $N_{2019} = 225$ ($N_{2020} = 311$) preservice teachers took part in the first online survey. Of those, $n_{2019} = 121$ ($n_{2020} = 119$) students took part in the survey at T5. Participation rates in surveys for each measurement occasion are depicted in Table 1. In

2019 participants were on average $M = 21.44$ ($SD = 2.36$) years old (2020: $M = 21.55$, $SD = 2.86$), and 66.2 % were female (2020: 64.6 %). 2019, students stated that they had been studying for $M = 4.21$ ($SD = 2.36$) semesters (2020: $M = 4.28$, $SD = 2.39$) and indicated studying a wide variety of teaching subjects.

Table 1

Number of Study Participants in Both Cohorts over the Course of the Semester.

Cohort	T1 *N*	T2 *n* (%)	T3 *n* (%)	T4 *n* (%)	T5 *n* (%)
2019	225	142 (63.1)	108 (47.9)	123 (54.6)	121 (53.7)
2020	311	192 (61.7)	162 (52.1)	145 (46.6)	119 (38.2)

Note. % in relation to T1 participants.

4.3 Measures

Next to the demographic data reported above (see 4.2) we assessed the following variables for our study:

Intrinsic Motivation for Educational Psychology (T1-T5) was assessed with three items adapted for the "educational psychology" context instead of "school" (e.g., "What I learn in educational psychology is interesting to me.") from the Scale for Assessing Subjective School-related Values (SESSW; Steinmayr & Spinath, 2010). Students were asked how each statement applied to them from 1 (*not at all*) to 5 (*very much*). Cronbach's α indicated good reliability of the scale at each measurement occasion (α = .85−.99).

Use of Learning Activities (T3 and T5) was assessed with three items, each concerning a learning activity that could be used during the semester: regularly submitting writing assignments, regularly attending the lecture sessions and taking self-test. For use of the activities *writing assignments* and *lecture attendance* students were asked, how much they had used them on a scale from 1 (*not at all*) to 5 (*very much*). For use of the activity *self-tests* students were asked the number of times they had participated in self-tests (1−5 at T5 and 1−3 at T3). For use of learning activities at T5 a mean was computed.

4.4 Data Analysis

All models were estimated with Mplus 8.0 (Muthén & Muthén, 1998-2017). Hypotheses were investigated using a multigroup second-order linear change model. Within this framework, intrinsic motivation at each time point was modelled as a latent state factor with three indicators each time. Further, second-order change factors (level and change) were estimated to load on latent state motivation factors. The level factor had loadings of one on all latent state factors of intrinsic motivation (T1−T5) and the change factor had loadings of 0, 0.3, 0.6, 0.9 and 1.2 on latent state factors of intrinsic motivation T1−T5 respectively because surveys took place every three weeks. This modeling approach accounts for measurement errors and bias in parameter estimates at the item and construct levels (Grimm & Ram, 2009). We evaluated model fit with the Comparative Fit Index (CFI; Bentler, 1990), the Root Mean Square Error of Approximation (RMSEA; Steiger, 1990) and the Standardized Root Mean Square Residual (SRMR; Kline, 2011). According to Hu and Bentler (1999), a CFI higher than .95 and RMSEA as well as SRMR lower than .08 indicate good fit.

Before applying the growth model, we tested for longitudinal measurement invariance to make sure we can compare factor means of intrinsic motivation across measurement occasions. To do so, strong measurement invariance is a prerequisite, which means that intercepts of manifest indicators as well as factor loadings need to be invariant over time. To test for strong invariance, we compared models of strong invariance with models with released constraints and evaluated if the latter showed meaningful better model fit. According to Chen (2007) a meaningful improvement of model fit when releasing intercepts is indicated by an increase of CFI $\geq .005$ and at the same time a decrease of RMSEA of $\geq .010$ or a decrease of SRMR $\geq .005$. When releasing factor loadings, an increase of CFI $\geq .005$ needs to be accompanied by a decrease of RMSEA of $\geq .010$ or a decrease of SRMR $\geq .025$.

To investigate motivational change, we estimated a multigroup linear change model for intrinsic motivation and inspected estimates of the change factors (H1). Next, we compared estimates for latent growth factors to evaluate differences between cohorts (H2) and entered students' mean activity use as a predictor of the change factor to investigate the association of evidence-based activities with motivational change (H3).

5. Results

Descriptive data for intrinsic motivation and uses of learning activities in both cohorts is depicted in Table 2.

Table 2

Means and Standard Deviations for Study Variables and Scales in Both Cohorts.

Variable	Cohort 2019 *M* (*SD*)	Cohort 2020 *M* (*SD*)
Intrinsic Value T1 [a]	3.98 (0.77)	3.95 (0.82)
Intrinsic Value T2 [a]	3.72 (0.88)	3.82 (0.81)
Intrinsic Value T3 [a]	3.77 (0.81)	3.79 (0.79)
Intrinsic Value T4 [a]	3.73 (0.85)	3.76 (0.75)
Intrinsic Value T5 [a]	3.66 (0.89)	3.71 (0.84)
Use of Learning Activities (T3)		
Lecture visit [b]	4.31 (1.14)	4.45 (1.06)
Writing assignments [b]	2.37 (1.47)	3.81 (1.34)
Number of self-tests [c]	2.12 (0.84)	2.14 (0.89)
Mean Use of Learning Activities (T5)	2.81 (1.29)	2.87 (1.52)
Lecture visit [b]	4.12 (1.35)	4.47 (1.11)
Writing assignments [b]	2.31 (1.34)	4.02 (1.11)
Number of self-tests [b]	3.21 (1.58)	2.98 (1.64)

Note. [a] Scale consisting of three items that could be answered on a scale ranging from 1 (*not at all*) to 5 (*very much*); [b] possible values from 1-5; [c] possible values from 1-3.

On a descriptive basis, students began the semester with relatively high intrinsic motivation. Although scores decreased over time, motivation remained high in both cohorts (see table 2 and figure 2).

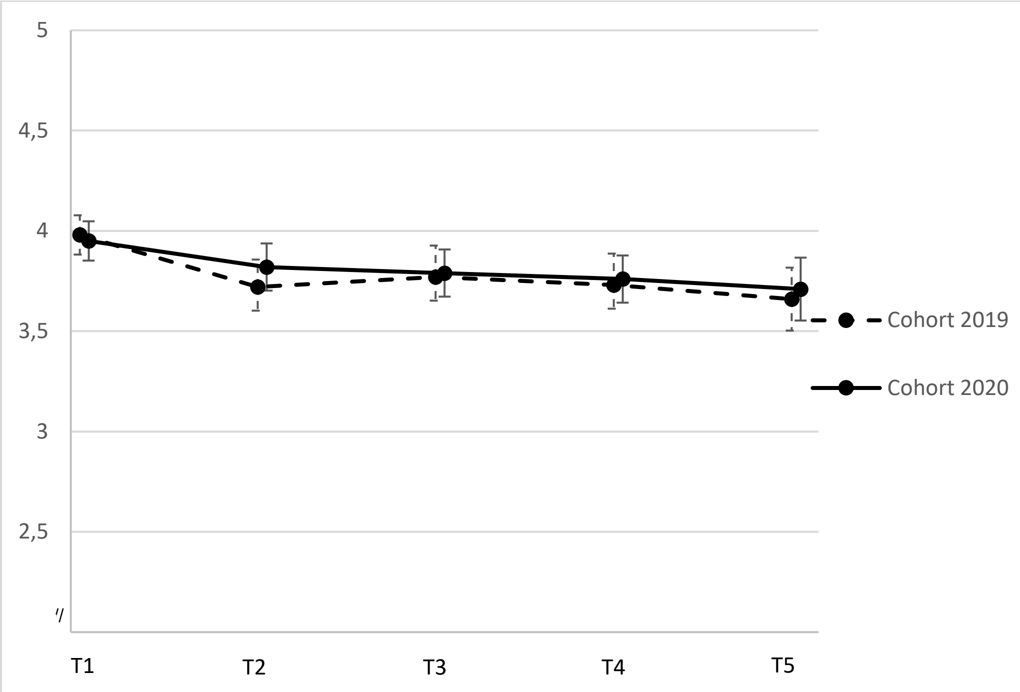


Figure 2. Means and 95 %-Confidence-Intervals of intrinsic motivation from t1-t5.

Preliminary analyses indicated strong measurement invariance over time. The model assuming strong invariance had a good model fit (CFI = .992; RMSEA = .024; SRMR = .040). Releasing invariance constraints of intercepts led to no meaningful improvement in model fit (CFI = .995; RMSEA = .020; SRMR = .036; according to Chen, 2007). Also, releasing invariance constraints of factor loadings led to no meaningful improvement in model fit (CFI = .995; RMSEA = .020; SRMR = .025). All further analyses were conducted with the model of strong measurement invariance.

The multigroup linear change model had an acceptable model fit (CFI = .965; RMSEA = .048; SRMR = .092). Standardized factor loadings were all above .74 (ps < .01) for observed motivation values loading on the corresponding latent motivation factor (T1−T5). Means of second-order change factors in both cohorts were negative and significantly different from 0 indicating a decline of intrinsic motivation. Further, variances of level as well as change factors in both cohorts were significantly different from 0 indicating interindividual differences in motivation at the beginning and in the motivational development over the course of the semester (see Table 3).

Table 3

Growth Factors of Intrinsic Motivation.

	Cohort 2019		Cohort 2020	
	Mean [CI]	*Var* [CI]	*Mean* [CI]	*Var* [CI]
Level	3.78 [3.66, 3.89]	0.42 [0.28, 0.56]	3.77 [3.67, 3.86]	0.42 [0.29, 0.55]
Change Factor	−0.29 [−0.43, −0.16]	0.34 [0.14, 0.54]	−0.32 [−0.44, −0.21]	0.19 [0.04, 0.35]

Note. Unstandardized Coefficients from Multigroup Linear Change Model; Var Variance; CI 95 % Confidence Interval.

Means' 95%-Confidence-Intervals of Level and Change factor between cohorts overlapped indicating no significant differences in motivation at the beginning or motivational development in 2020 compared to 2019.

When adding the mean of students' learning activity use over the course of the semester (t5) as a manifest predictor of the Change factor, controlling for its' correlation with the Level factor at the same time, model fit remained acceptable (CFI = .964; RMSEA = .045; SRMR = .106). Students' mean use of all learning activities was positively associated with change of intrinsic motivation in both cohorts ($r_{2019} = .489$, $p < .01$; $r_{2020} = .453$, $p < .01$). Looking at lecture attendance on the one hand and additional activities (self-tests and writing assignments) on the other hand, the following results were obtained: The use of additional activities was associated with motivational change in 2020 ($r_{2020} = .411$, $p < .01$) but not in 2019 ($r_{2019} = .232$, $p = .118$). However, constraining the correlation coefficients to be equal across groups did not change model fit significantly ($\Delta\chi^2$ (1) = 0.049, $p = .825$), indicating that the coefficients between groups did not differ significantly. When looking only at lecture attendance, this activity explained variance in motivational development in a face to face semester ($r_{2019} = .567$, $p = .017$) but not in an online semester ($r_{2020} = .051$, $p = .704$).

Constraining these coefficients to be equal across groups led to a significantly worse model fit ($\Delta\chi^2$ (1) = 6.485, p = .011) indicating that the coefficients between groups differed significantly.

T-tests with an adjusted $\alpha < .01$ (Bonferroni corrected) were computed to compare students who dropped out before T5 and who completed all surveys. We neither found differences in intrinsic motivation (T1-T4) in either cohort (t(106-309) = −0.213−2.092, all ps > .01) nor differences in class attendance at T3 (2019: t(106) = 0.485, p > .01; 2020: t(159) = 2.168, p > .01). However, students who dropped out until T5 already reported significantly less submissions of writing assignments at T3 in cohort 2020 (t(158) = 4.507, p < .01, d = 0.81, 95 % CI [0.469, 1.148]) and significantly less participation in self-tests both in cohort 2019 (t(226) = 8.850, p < .01, d = 1.17, 95 % CI [0.892, 1.455]) and in cohort 2020, t(309) = 12.048 , p < .01, d = 1.41, 95 % CI [1.151, 1.659].

6. Discussion and Conclusion

Different from what was expected in the face of the COVID pandemic and possible motivational challenges, students' initial intrinsic motivation and change of intrinsic motivation did not differ between a face-to-face and an online semester. Surprisingly, in both cohorts, students' motivation was high at the beginning of the semester and stayed high, instead of the overall decline. Other authors have reported detrimental changes in students' motivation (Müller et al., 2021). However, the authors investigated the decreased possibilities of need satisfaction (e.g., peer interaction to feel related) relying on Ryan and Deci's (2020) self-determination approach. Need satisfaction can be considered a prerequisite for intrinsic motivation. However, intrinsic motivation can be explained by more factors and the learners in our study maybe were able to cope with motivational challenges. In addition, the study by Usher and colleagues (2021) retrospectively asked students about motivational changes. Considering that other variables such as student well-being have suffered (e.g., Schwinger et al., 2020; Steinmayr et al., 2022) students' retrospective data on motivational change may be biased. In our study we were able to compare students of the same lecture class from 2019 and 2020 with the same measures of intrinsic motivation. The sample in both cohorts were comparable regarding their sociodemographic characteristics and lecture content did not differ between cohorts. This allowed us to compare the change in motivation directly within one model and we could not find motivational differences between cohorts. One explanation could be that students who experienced a sharp drop in motivation right at the beginning in 2020 dropped out of the data collection. However, because motivation did not differ between cohorts at baseline, other explanations are possible. Schwinger and colleagues (2020) reported that it was primarily the decrease in autonomy during the lockdown that led to a decrease in well-being. It is possible that our students experienced no loss of autonomy within the lecture and were able to focus on lecture content and engage in more learning activities because constraints such as social distancing reduced other activities and sources of satisfaction. Another possible explanation for the comparable motivational development lies in the evidence-based activities themselves. Perhaps these activities made distance learning more attractive in the context of this lecture, which would highlight the opportunities of evidence-based learning activities is distance education.

Students in both cohorts seemed to suffer from motivational challenges reflected in a motivational decline and dropout. Recently, Benden and Lauermann (2021) highlighted the importance of short-term motivational changes for performance and course dropout in university. The motivational decline we found, highlights a challenge for university students and instructors that is present both in face-to-face courses and in distance education. When interpreting the decline in motivation, it should be considered that motivation was quite high at the beginning of the semester. In addition, almost everyone probably knows the feeling that working on a long-term project - whether it is successfully attending a lecture as a student, teaching a lecture as an instructor, or working on any project - challenges motivation from time to time. Motivational problems do not mean that one is certain to fail. How one deals with challenges is critical.

To encounter students' motivational challenges evidence-based learning activities may be a suitable tool. Despite from being helpful for students' performance (Bosch et al., 2021), in this study they were related to a more positive development of motivation. The associations of learning activities and motivational change were moderate to large in size. Of course, motivation is also important to engage in learning activities. We argue, however, that the continued engagement may also help to maintain motivation over time because learning behavior is not only a consequence, but also a prerequisite as many models of motivation include past behavior as a precursor of motivation (e.g., Eccles & Wigfield, 2020). Evidence-based learning activities can be considered a chance for university instructors and students, because they could be easily adapted to different contexts and are associated to students' motivation even in pandemic-related distance education. In both cohorts, students could attend the lecture (in presence in summer 2019; online in summer 2020) to get to know the content and to ask questions. This activity probably differed the most between 2019 and 2020. Explorative analyses revealed that in 2020 additional evidence-based activities (writing assignments and self-tests) were associated with a more favorable development of motivation. Given the reduced interaction between students and instructors and less feedback in distance education, individual feedback on tests as well as the more structuring character of the activities (particular time slots for each activity) might have gained importance for students' motivation.

Overall, dropout from surveys and also from the course was high, especially in 2020. One reason for the higher dropout rate 2020 could be that more students visited the first lecture session and enrolled in the course (N = 311 compared to N = 225 in 2019) because it was easier online and they may did not have as strong intentions to complete the course. Further, students who submitted more writing assignments and participated in more self-tests completed all surveys and probably the course with a higher probability. This points to another important aspect of evidence-based learning activities. They do help students learn the content, they are related to a more favorable motivational development and they may also help students finish a course (maybe by preventing them from drastic motivational declines).

Our study is the first that compared student motivation and its' development over the course of one semester in a face-to-face lecture course and immediately after the change to distance education and associations with evidence-based learning activities. There are, of course, limitations that need to be considered. First of all, we cannot draw causal conclusions about the effect of learning activities due to the non-experimental design. However, applying a longitudinal design gives us first ideas how motivational challenges in higher education could be encountered. Further, it is likely that our findings are limited to certain contextual conditions. Our study was conducted in a large lecture class with specific evidence-based learning activities. In lecture classes, even in a face-to-face semester, social interaction is usually not as lively as in seminar formats. Therefore, the discrepancy in social interaction and maybe also motivation might have been more pronounced in seminar formats. However, the study of large lecture classes in distance education and the ways in which evidence-based activities can be implemented here are of great importance to educational practice, as large lecture classes are a common format in undergraduate curricula. The generalizability of these findings to other student groups and the further development of motivational trends in ongoing distance education are beyond the scope of this study and would be of interest for future research.

In addition, we had to rely on self-report data for the study variables. The assessment of learning activities might be better if they were collected objectively (e.g., the number of essays a student submitted). In this study, we had to rely on self-reports because of anonymous data collection. However, we asked about specific activities, the survey was voluntary and anonymous, and students did not face any consequences in case they did not participate, so we assume that they answered truthfully.

Considering dropout rates, our results could be biased because more motivated students participated in more surveys. However, we still had some variability in our sample regarding motivation and our analyses indicated that students who participated in all surveys did not differ in motivation from those who dropped out.

With these limitations in mind, our study yielded noteworthy results. In contrast to the results of other empirical studies, student motivation does not seem to have suffered more in our case with distance learning than with face-to-face learning. One aim was to explore chances to deal with motivational challenges that students experience over the course of one semester. In the case of distance learning as well as face-to-face learning, we highlighted the opportunities for teachers to provide useful and potentially motivating activities. This could help motivate students by keeping them engaged continuously.

References

Balgopal, M.M., Wallace, A.M., & Dahlberg, S. (2012). Writing to learn ecology: A study of three populations of college students. *Environmental Education Research*, *18*(1), 67–90. https://doi.org/10.1080/13504622.2011.576316

Benden, D.K., & Lauermann, F. (2021). Students' motivational trajectories and academic success in math-intensive study programs: Why short-term motivational assessments matter. *Journal of Educational Psychology*. Advance online publication. https://doi.org/10.1037/edu0000708

Bentler, P.M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107,* 238–246. https://doi.org/10.1037/0033-2909.107.2.238

Bosch, E., Seifried, E., & Spinath, B. (2021). What successful students do: Evidence-based learning activities matter for students' performance in higher education beyond prior knowledge, motivation, and prior achievement. *Learning and Individual Differences*, *91*, 102056. https://doi.org/10.1016/j.lindif.2021.102056.

Carpenter, S.K. (2012). Testing enhances the transfer of learning. *Current Directions in Psychological Science*, *21*(5), 279–283. https://doi.org/10.1177/0963721412452728

Chen, F.F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14,* 464–504. https://dx.doi.org/10.1080/10705510701301834

Cogliano, M., Kardash, C.M., & Bernacki, M.L. (2019). The effects of retrieval practice and prior topic knowledge on test performance and confidence judgments. *Contemporary Educational Psychology*, *56*, 117–129. https://doi.org/10.1016/j.cedpsych.2018.12.001

Credé, M., & Kuncel, N.R. (2008). Study habits, skills, and attitudes: The third pillar supporting collegiate academic performance. *Perspectives on Psychological Science*, *3*(6), 425–453. https://doi.org/10.1111/j.1745-6924.2008.00089.x

Credé, M., Roch, S.G., & Kieszczynka, U.M. (2010). Class attendance in college: A meta-analytic review of the relationship of class attendance with grades and student characteristics. *Review of Educational Research*, *8*(2), 272–295. https://doi.org/10.3102/0034654310362998

Downs, S.D. (2015). Testing in the college classroom: Do testing and feedback influence grades throughout an entire semester? *Scholarship of Teaching and Learning in Psychology, 1*(2), 172–181. https://doi.org/10.1037/stl0000025

Dunlosky, J., Rawson, K.A., Marsh, E.J., Nathan, M.J., & Willingham, D.T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, *14*(1), 4–58. https://doi.org/10.1177/1529100612453266

Dunn, D.S., Saville, B.K., Baker, S.C., & Marek, P. (2013). Evidence-based teaching: Tools and techniques that promote learning in the psychology classroom. *Australian Journal of Psychology*, *65*(1), 5–13. https://doi.org/10.1111/ajpy.12004

Dutke, S., & Barenberg, J. (2015). Easy and informative: Using confidence-weighted true–false items for knowledge tests in psychology courses. *Psychology Learning & Teaching*, *14*(3), 250–259. https://doi.org/10.1177/1475725715605627

Eccles, J.S., Adler, T.F., Futterman, R., Goff, S.B., Kaczala, C.M., Meece, J.L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives* (pp. 75–146). San Francisco: Freeman.

Eccles, J.S., & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology*, *61*, 101859. https://doi.org/10.1016/j.cedpsych.2020.101859

Garris, C.P., & Fleck, B. (2022). Student evaluations of transitioned-online courses during the COVID-19 pandemic. *Scholarship of Teaching and Learning in Psychology*, *8*(2), 119-139. https://doi.org/10.1037/stl0000229

Grimm, K.J., & Ram, N. (2009). Nonlinear growth models in M plus and SAS. *Structural Equation Modeling*, *16*(4), 676–701. https://doi.org/10.1080/10705510903206055

Heublein, U., & Wolter, A. (2011). Studienabbruch in Deutschland. Definition, Häufigkeit, Ursachen, Maßnahmen [Dropout from higher education in Germany - Definition, dropout rate, causes, measures]. *Zeitschrift für Pädagogik*, *57*(2), 214–236. https://doi.org/ 10.25656/01:8716

Hidi, S., & Renninger, K.A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*(2), 111–127. https://doi.org/10.1207/s15326985ep4102_4

Hu, L., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1–55. http://dx.doi.org/10.1080/10705519909540118

Jacobs, J.E., Lanza, S., Osgood, D.W., Eccles, J.S., & Wigfield, A. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. *Child Development*, *73*(2), 509–527. https://doi.org/10.1111/1467-8624.00421

Jansen, T., Meyer, J., Wigfield, A., & Möller, J. (2022). Which student and instructional variables are most strongly related to academic motivation in K-12 education? A systematic review of meta-analyses. *Psychological Bulletin*, *148*(1-2), 1–26. https://doi.org/10.1037/bul0000354

Kedraka, K., and Kaltsidis, C. (2020). Effects of the COVID-19 pandemic on university pedagogy. Students experiences and considerations. *European Journal of Educational Studies, 7*(8), 17–30. https://doi.org/10.46827/ejes.v7i8.3176

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

Kosovich, J.J., Flake, J.K., & Hulleman, C.S. (2017). Short-term motivation trajectories: A parallel process model of expectancy-value. *Contemporary Educational Psychology, 49*, 130–139. https://doi.org/10.1016/j.cedpsych.2017.01.004

Kriegbaum, K., Becker, N., & Spinath, B. (2018). The relative importance of intelligence and motivation as predictors of school achievement: A meta-analysis. *Educational Research Review*, *25*, 120–148. https://doi.org/10.1016/j.edurev.2018.10.001

Lockl, K., Attig, M., Nusser, L., & Wolter, I. (2021). Cognitive and Affective-Motivational Factors as Predictors of Students' Home Learning During the School Lockdown. *Frontiers in Psychology*, *12*, 751120–751120. https://doi.org/ 10.3389/fpsyg.2021.751120

Muthén, L.K., & Muthén, B.O. (1998 –2017). *Mplus user's guide* (8th ed.). Los Angeles, CA: Author.

Müller, F.H., Thomas, A.E., Carmignola, M., Dittrich, A.K., Eckes, A., Großmann, N., ... & Bieg, S. (2021). University students' basic psychological needs, motivation, and vitality before and during COVID-19: A Self-Determination Theory approach. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.775804

OECD (2013). *Education at a glance: OECD indicators*. Paris: OECD.

Phelps, R.P. (2012). The effects of testing on student achievement, 1910–2010. *International Journal of Testing, 12,* 21– 43. https://doi.org/10.1080/15305058.2011.602920

Pintrich, P.R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review*, *16*(4), 385–407. https://doi.org/1040-726X/04/1200-0385/0

Robinson, K.A., Lee, Y.K., Bovee, E.A., Perez, T., Walton, S.P., Briedis, D., & Linnenbrink-Garcia, L. (2019). Motivation in transition: Development and roles of expectancy, task values, and costs in early college engineering. *Journal of Educational Psychology*, *111*(6), 1081–102. https://doi.org/10.1037/edu0000331

Ryan, R.M., & Deci, E.L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology*, *61*, 101860. https://doi.org/10.1016/j.cedpsych.2020.101860

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. https://doi.org/10.1016/j.cedpsych.2005.02.002

Schwinger, M., Trautner, M., Karchner, H., & Otterpohl, N. (2020). Psychological impact of Corona

lockdown in Germany: Changes in need satisfaction, well-being, anxiety, and depression. *International Journal of Environmental Research and Public Health, 17*(23), 9083. https://doi.org/10.3390/ijerph17239083

Seifried, E., & Spinath, B. (2020). Using formative self-assessment to improve teaching and learning in educational psychology courses. In E. Landrum, C. Hakala, & S. Nolan (Eds.), *Assessing Undergraduate Learning in Psychology: Strategies for Measuring and Improving Student Performance.* APA Books.

Steiger, J.H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research, 25,* 173–180. http://dx.doi.org/10.1207/s15327906mbr2502_4

Steinmayr, R., Paschke, P., & Wirthwein, L. (2022). Elementary School Students' Subjective Well-Being Before and During the COVID-19 Pandemic: A Longitudinal Study. *Journal of Happiness Studies*, 1–21. https://doi.org/10.1007/s10902-022-00537-y

Steinmayr, R., & Spinath, B. (2010). Konstruktion und erste Validierung einer Skala zur Erfassung subjektiver schulischer Werte (SESSW) [Construction and First Validation of a Scale for Assessing Subjective School-related Values]. *Diagnostica, 56*, 195–211. https://doi.org/10.1026/0012-1924/a000023

Trautwein, U., Marsh, H.W., Nagengast, B., Ludtke, O., Nagy, G., & Jonkmann, K. (2012). Probing for the multiplicative term in modern expectancy-value theory: A latent interaction modeling study. *Journal of Educational Psychology, 104*(3), 763–777. https://doi.org/10.1037/a0027470.

Usher, E.L., Golding, J.M., Han, J., Griffiths, C.S., McGavran, M.B., Brown, C.S., & Sheehan, E.A. (2021). Psychology students' motivation and learning in response to the shift to remote instruction during COVID-19. *Scholarship of Teaching and Learning in Psychology*. Advance online publication. https://doi.org/10.1037/stl0000256

Vansteenkiste, M., Ryan, R.M., & Soenens, B. (2020). Basic psychological need theory: Advancements, critical themes, and future directions. *Motivation and Emotion*, *44*(1), 1-31. https://doi.org/10.1007/s11031-019-09818-1

Wang, M.T., & Eccles, J.S. (2012). Social support matters: Longitudinal effects of social support on three dimensions of school engagement from middle to high school. *Child Development, 22*, 31–39. https://doi.org/10.1111/j.1467-8624.2012.01745.x

Wigfield, A., & Eccles, J.S. (2000). Expectancy - value theory of motivation. *Contemporary Educational Psychology, 25*(1), 68–81. https://doi.org/10.1006/ceps.1999.1015

Wong, R. (2020). When no one can go to school. Does online learning meet students' basic learning needs? *Interactive Learning Environments* 72, 1–17. https://doi.org/10.1080/10494820.2020.1789672

Paper III

This is the accepted version of the article

Bosch, Eva, & Spinath, Birgit (2023). What evidence-based learning activities help students acquire knowledge, correct confidence in their own knowledge, and accurate self-assessment? *Learning and Individual Differences, 108*, 102374. https://doi.org/10.1016/j.lindif.2023.102374

CRediT author statement

Eva Bosch: conceptualization, investigation, data curation, formal analysis, writing – original draft.

Birgit Spinath: conceptualization, supervision, investigation, writing – review & editing.

What evidence-based learning activities help students acquire knowledge, correct confidence in their own knowledge, and accurate self-assessment?

Eva Bosch[1*], & Birgit Spinath[1]

[*]corresponding author; Eva.Bosch@psychologie.uni-heidelberg.de

[1]Heidelberg University, Department of Psychology, Hauptstraße 47-51, 69117 Heidelberg, Germany

Abstract

The current study examines the development of higher education students' knowledge, correct confidence in their own knowledge, and accuracy of self-assessment over one semester and if the use of evidence-based learning activities (such as practice testing) is associated with these outcomes. The present study assessed $N = 285$ university students' use of learning activities, their knowledge, correct confidence, and accuracy of self-assessment five times over the course of a semester in a lecture course. All learning outcomes improved and were explained by learning activity use beyond students' prior knowledge, prior achievement and motivation. Latent profile analyses distinguished five subgroups of students regarding the frequency and combination of learning activity use. Results indicate that students benefitted most from using multiple activities. The study provides evidence that in a realistic learning setting several learning activities contribute to students' knowledge, correct confidence, and accuracy of self-assessment gains.

1. Introduction

In higher education, students need and want to acquire knowledge to pass exams and continue their studies. Models of self-regulated learning hold that metacognitive monitoring is essential for students' learning success (Schmitz & Wiese, 2006; Winne, 1996; Zimmerman, 2008). When students metacognitively monitor their learning, they can assess their own knowledge and adjust their learning accordingly. Indeed, empirical studies, have shown that accurate self-assessment of one's knowledge helps students to successfully regulate their learning and is associated with better performance (Thiede, 1999; Thiede et al, 2003).

Evidence-based learning activities might help students to gain knowledge as well as an accurate self-assessment of their own knowledge. Various learning activities that can be provided in higher education have been shown to explain students' knowledge acquisition (e.g., Bae et al., 2019; Dunlosky et al., 2013). Self-tests in particular have also been shown to help students accurately assess their own knowledge (e.g., Barenberg & Dutke, 2019; Naujoks et al., 2022). However, testing is not the only learning activity, and in real learning situations, it is likely that students employ various activities, potentially in individually varying combinations. To investigate which learning activities students use and combine, several optional activities were offered to a cohort of university students in this study. The use of the activities and possible combinations were then evaluated with a person-centered approach. We investigated how the use of different evidence-based learning activities contributes not only to knowledge acquisition but also to students' metacognitive assessment of their own knowledge.

2. Self-regulation and supply-use models for successful learning in higher education

To understand successful learning in higher education, we build on process models of self-regulated learning (e.g., Schmitz & Wiese, 2006; Winne, 1996; Zimmerman, 2008) as well as supply-use models of learning (Bosch et al., 2021; Brühwiler & Blatchford, 2011; Seidel, 2014).

The learning process in higher education is characterized by little external control of students' progress and engagement during a semester and by several opportunities for students to choose when and how to study, i.e., students learn self-regulated. In models of self-regulated learning, learning is viewed as a cyclical process in which students continuously plan, act, and evaluate (Schmitz & Wiese, 2006; Winne, 1996; Zimmerman, 2008). By evaluating their own learning process and outcomes, learners create an internal feedback loop that is crucial for self-regulation. After a learning action (postaction phase; Schmitz & Wiese, 2006), learners evaluate the process and outcomes of the action phase in order to plan the next learning activity. To evaluate correctly, learners must have an accurate assessment of their learning process and outcomes. The relationship between actual learning and learners' perceptions of learning is referred to as metacognitive monitoring (Winne, 1996; Zimmerman, 2008) and is assessed by comparing the actual learning process and outcomes with learners' perceptions of it (e.g., Barenberg & Dutke, 2019; Hacker et al., 2000). When subjective perceptions of learning match actual learning and metacognitive monitoring is accurate, students can identify errors and set appropriate learning goals for the next learning activity (forethought/planning; Schmitz & Wiese, 2006; Winne, 1996; Zimmerman, 2008).

In this study, we investigate two facets of metacognitive monitoring. First, the comparison between students' confidence in their knowledge of specific topics and their actual knowledge of those specific topics, i.e., students' awareness of content-specific strengths and weaknesses (hereafter, correct confidence; see also Barenberg & Dutke, 2019). Second, we consider the comparison between students' perceived level of their knowledge and their actual level of knowledge across all content (hereafter, accuracy of self-assessment; see also Hacker et al., 2000).

Supply-use models of learning consider that learning is always an interplay between instructors who provide learning opportunities and learners who use them (e.g., Brühwiler & Blatchford, 2011). In higher education, a common learning situation in different fields of study is lectures with weekly sessions taught by a lecturer and an exam at the end of a course (Rutkiene & Tandzegolskiene, 2015; Watts & Schaur, 2011). When optional evidence-based learning activities are offered to prepare for such an exam, students must decide whether to use them, which ones to use and to what extent, and act accordingly, oftentimes without external control. Successfully planning and deciding on learning activities requires students to carefully consider their current level of knowledge, their own content specific strengths and weaknesses, and to evaluate given learning opportunities. By combining requirements of self-regulated learning with the supply-use nature of learning situations in higher education, we aim to obtain a detailed picture of students' learning behaviors as well as their learning outcomes.

2.1 Indicators of learning success: knowledge, correct confidence, accuracy

Learning success in higher education is typically examined through test scores or performance on exams that assess the knowledge students have acquired (Credé et al., 2010; Richardson et al., 2012; Robbins et al., 2004; Sackett et al., 2009; Schneider & Preckel, 2017). Although more knowledge is undeniably an important goal of higher education teaching, students should not only acquire knowledge, but also gain correct confidence and become aware of how much they know during the learning process.

In self-regulated learning situations, students are thought to make many decisions regarding their learning (Schmitz & Wiese, 2006; Winne, 2011; Zimmerman, 2008). These decisions accompany the entire learning process: when, where and how to learn, which learning activity to use, which content to learn or prioritize, but also when to stop learning. In order to decide correctly, what content to focus on, students need to metacognitively monitor what they already know and what there is still to learn (see section 2. on models of self-regulated learning by Winne, 1996; Zimmerman, 2008). Whether students correctly assess their own content-specific strengths and weaknesses can be examined in a knowledge test using item- or content-specific confidence ratings and the resulting precision of confidence, i.e., whether the confidence and correctness of the answers match (correct confidence; see also Barenberg &

Dutke, 2019). To decide how much time to invest in studying or when to stop studying, students need to accurately assess how much of the relevant content they already know (Winne, 1996; Zimmerman, 2008). Whether a student accurately assesses his or her level of knowledge can be examined by comparing the student's actual performance (test scores) and self-assessment of his or her performance (estimated test scores; see also Foster et al., 2017; Naujoks et al., 2022).

According to the models of self-regulated learning described in section 2. both correct confidence and accuracy of self-assessment should enable students to use their study time for less familiar content and realistically decide whether to continue learning (Thiede, 1999; Winne, 1996). In line with these considerations, empirical research has shown that correct confidence in one's own knowledge as well as an accurate estimate of one's own knowledge level help students to successfully learn self-regulated (Händel et al., 2020; Kostons et al., 2012; Stone, 2000; Thiede et al., 2003). For this reason, we consider knowledge acquisition but also correct confidence and accuracy of self-assessment indicators of successful learning. However, students are often overconfident (e.g., Bensley et al., 2015; Händel & Dresel, 2018) and tend to misjudge their level of knowledge (Bensley et al., 2015; Dunning et al., 2003; Foster et al., 2017; Hartwig & Dunlosky, 2014; Miller & Geraci, 2011). Therefore, we are interested how knowledge, correct confidence, and accuracy of self-assessment change over the course of a semester when students attend a lecture class and if these learning outcomes are explained not only by individual learning prerequisites (such as prior knowledge, prior achievement and motivation) but also the use of various learning activities.

2.2 Explaining learning success: importance of evidence-based learning activities

A large body of empirical research has identified prior knowledge, prior achievement, motivation as well as study skills and habits as important predictors of learning success in higher education (for meta-analyses and reviews see e.g., Credé & Kuncel, 2008; Kriegbaum et al., 2018; Richardson et al., 2012; Schneider & Preckel, 2017). Therefore, the current study considers important individual learning prerequisites, i.e., prior achievement, prior knowledge, and motivation in order to replicate their associations with students' knowledge acquisition and also to examine additional benefits of specific learning activities.

2.2.1 Practice testing

Research on learning habits and specific learning activities has shown that various learning activities contribute to students' learning success beyond learning prerequisites (e.g., Bosch et al., 2021; Credé & Kuncel, 2008). Within evidence-based teaching, learning activities should be implemented that have shown to promote student learning in controlled settings (see recommendations by Boser et al., 2017). A frequently investigated learning activity is practice testing. Practice testing is a highly useful learning activity to improve performance (Dunlosky et al., 2013). A recent meta-analytic review of over 200 classroom studies by Yang and colleagues (2021) reported beneficial effects of practice testing on students' performance in comparison to other learning activities (g = 0.49) also for untested knowledge, confirming the usefulness of tests as a learning activity for exam preparation (g = 0.32). This *testing-effect* might be explained by several mechanisms, including increased retrieval effort, which is thought to enhance learning (Rowland, 2014), benefits of similarity between practice and assessment (Thomas & McDaniel, 2013), and increased motivation to prepare for tests (Yang et al., 2021).

Adding to the benefits of testing for knowledge acquisition, empirical studies have also found that testing is associated with indicators of metacognitive monitoring (e.g., Barenberg & Dutke, 2019; Carvalho et al., 2022; Naujoks et al., 2022). Barenberg and Dutke (2019) argue that students could use information from earlier retrieval attempts as a cue to judge the outcome of the current test. Further, retrieval might stimulate the elaboration of the tested content thereby increasing accessibility of relevant information and facilitating future confidence ratings. In this line, several empirical studies have found an association between testing and correct confidence. For example, Barenberg and Dutke (2019) found secondary school students' correct confidence to be greater in a final test when learning with retrieval tests (d = 0.49). In a quasi-experimental study, Händel and colleagues (2020) had undergraduate students answer multiple-choice questions in a final test. Students were then asked to provide item-specific judgements of their answers, indicating their awareness of content-specific strengths and weaknesses (correct confidence). Students in the testing group exhibited significantly higher correct confidence than those in the control condition, with a medium to large effect size ($\eta^2 = 0.14$). Naujoks and colleagues (2022) allowed undergraduate students to self-test voluntarily up to four times during one semester. They evaluated students' item-specific judgements in the final exam and observed that increased practice testing correlated with higher correct confidence on the final course exam ($R^2 = .03$).

While students' knowledge and correct confidence seem to benefit from testing, empirical research that focused on self-assessment of own knowledge-level (accuracy) has revealed mixed evidence (e.g., Fernandez & Jamet, 2016; Foster et al., 2017; Rivers et al., 2019). For example, undergraduate students in Foster and colleagues' (2017) study were unable to improve their accuracy in predicting their test score even after 13 practice tests with feedback, whereas undergraduate students in Fernandez and Jamet's (2016) testing condition provided a significantly more accurate prediction of their final exam grade than students from a control condition. Whether accuracy of self-assessment increases with testing, is, to date, an open question.

2.2.2 Designing a lecture course with multiple evidence-based learning activities

Despite the large number of studies on the effects of practice testing, in a recent study Carvalho and colleagues (2022) note that "real-world evidence for the benefits of practice testing is, at best, mixed, and potentially lacking. Moreover, although there has been some initial investigation […] in natural contexts […] these inquiries have been limited in scope both in terms of duration of practice (e.g., only 14 days) and type of materials (e.g., only multiple-choice testing questions)." (p. 1724). Further, studies on practice testing seldom consider practice testing as one activity among others. In this line, Naujoks and colleagues (2022) critically discuss that in real learning situations, further learning activities could influence the learning process over the course of the semester, and thus perhaps correct confidence and accuracy of self-assessment. Repeated exposure to and reflection on the content, for example during writing assignments, review of lecture notes, participation in lecture sessions and discussions, could stimulate elaboration of the content and increase familiarity with the content, so that correct confidence and accuracy of self-assessment could increase. However, to our knowledge no study to date has examined the associations of multiple learning activities in a natural learning situation with multiple learning outcomes, namely knowledge, correct confidence, and accuracy of self-assessment. In the current study, we therefore investigate whether use of several evidence-based learning activities contributes to these learning outcomes over the course of one semester.

According to the supply-use model for learning in higher education (Bosch et al., 2021), instructors can offer several evidence-based activities to students. Studies investigating students' actual learning behavior in higher education courses have highlighted students' tendency to cram shortly ahead of an exam and using a mixture of effective and ineffective study habits (e.g., Blasiman et al., 2017). Therefore, learning activities should be offered continuously and guide students to space their learning. When students are guided to space their learning with a continuous supply of regular optional learning activities over the course of one semester, this might help them not to cram shortly ahead of the exam and could enhance knowledge as well as correct confidence and accuracy (Barenberg et al., 2018; Dunlosky et al., 2013; Händel et al., 2020; see also Naujoks et al., 2022). Further, when students receive feedback on learning activities this could improve learning outcomes (for benefits of feedback, see, e.g., Downs, 2015; Hattie, 2011; Yang et al., 2021).

The activities offered and investigated in the present study are summarized below. In the present study, testing is investigated as one optional learning activity among others. Knowledge tests are designed to prepare students for the final exam. They include confidence-weighted true/false items that cover content from the last lectures and correspond to the format of the final exam. The second learning activity consists of the opportunity to submit essays answering deepening questions on the last lecture. Such writing exercises have been shown to be related to performance. For example, in a recent meta-analytic review, writing was found to improve learning with a medium effect size (Hedges g = .41; Schindler & Richter, 2023). Students receive feedback on these two activities to help them improve their performance (Downs, 2015; Yang et al., 2021). Next to these two learning activities two further activities were implemented and investigated. Possible regular and also common learning activities in lecture classes are the participation in lecture sessions and review of lecture slides and notes. Attending lectures and regularly reviewing slides and notes should help students acquire knowledge as it provides them with relevant information, exposure to the content in a variety of ways, but also allows them to distribute their practice (Carrier, 2003; Credé et al., 2010). In a large meta-analytic review, Credé and colleagues (2010) found that regular lecture attendance explained an additional 19 % of the variance in college students' performance beyond their prior achievement. Especially when lecture attendance is combined with independent review of the material, students distribute their practice and improve their performance (Cepeda et al., 2006; see also Credé et al., 2010). Lecture attendance and review of lecture notes could also improve correct confidence and accuracy of self-assessment because students become more familiar with the content and gain a more elaborated knowledge base as well as increased accessibility of relevant information.

If students engage in these diverse activities they might improve their knowledge, correct confidence, and accuracy of self-assessment. It is reasonable to assume that participants from all kinds of classroom-studies participated in various learning activities besides the ones included in the study design. By assessing, for example the frequency of class attendance in addition to testing activities in a real learning situation, we aim to obtain a realistic and detailed understanding of the benefits of students' learning behavior.

Considering the diversity of the learning activities class attendance, review of slides and notes, knowledge tests, and essays, regarding their difficulty, time-intensity and learning focus students may prefer some activities over others. Previous research has shown that students engage in multiple learning activities in self-regulated learning situations, at least some of which are relatively ineffective (e.g., Kornell & Bjork, 2007), although students seem to have fairly good knowledge of which learning activities are actually effective (Blasiman et al., 2017). Consequentially, we cannot assume that all students would use all offered evidence-based activities to a similar amount. In line with assumptions of self-regulated learning, it might be adaptive to focus on some learning activities that address individual learning needs and preferences, for example, study alone vs. attend classes. If we only look at the total amount of all activities used, we may miss some specific patterns of activity use and their effectiveness for some students. To consider this, we complement our analyses with a person-centered approach. Person-centered approaches such as LPA consider possible intra-individual variation within a system of variables (Marsh et al., 2009). Rather than focusing on one learning activity per se and how it relates to learning outcomes in the whole population, LPA allows us to identify subgroups of individuals sharing a similar pattern of learning activity use. In a next step, these subgroups can be compared regarding different learning outcomes.

3. Summary and research questions

Building on models of self-regulated learning (Schmitz & Wiese, 2006; Zimmerman, 2008) and a supply use model of learning in higher education (Bosch et al., 2021) we investigate students' learning behavior and outcomes in a real learning situation. With the current study we aim to answer three superordinate research questions. First, we want to examine several learning outcomes over the course of one semester in an educational psychology lecture course at university. Second, we want to investigate whether the use of evidence-based learning activities explains students' learning outcomes beyond their learning prerequisites.

RQ1) How do students' knowledge, correct confidence, and accuracy change over the course of a semester? We expect that students'

H1a: knowledge

H1b: correct confidence in their own knowledge and

H1c: accuracy of self-assessment

increase over the course of one semester.

RQ2) Does learning activity use explain learning outcomes? Students' learning activity use explained performance in a knowledge test in a prior study (Bosch et al.,, 2021). We wanted to replicate and expand this finding. In detail, we expect that students' use of learning activities over the course of one semester explains

H2a: knowledge

H2b: correct confidence in their own knowledge and

H2c: accuracy of self-assessment

at the end of the semester beyond students' learning prerequisites, i.e., prior knowledge, prior achievement, and motivation.

Last but not least we would like to extend these findings and investigate students' actual learning activity use in a real learning situation, identify possible patterns of use and related outcomes.

RQ3) Are there subgroups of students that use some learning activities more while using others less? Which specific learning activity profiles can be identified?

We had no specific expectations regarding particular learning activity profiles. However, as mentioned earlier, given the variety of activities, we expected that at least some students would employ patterns of activity use that were not just a more or less of all activities. For example, some students may decide on only one activity that seems most important to them and skip all other activities (e.g., only review lecture notes as part of a "study alone and schedule-independent" strategy).

RQ4) Are certain learning activity profiles associated with more knowledge, correct confidence, and accuracy of self-assessment? Which learning activities are associated with which outcomes?

4. Method

To address the research questions and test the hypotheses stated above, self-report data as well as performance in knowledge tests from undergraduate students were obtained multiple times during an academic semester.

4.1 Procedure

Undergraduate students attending introductory lecture classes on educational psychology were surveyed online five times over the course of one semester. The lecture classes were offered to two cohorts of students, undergraduate psychology students and undergraduate students from various subjects studying with an option to become teachers (preservice teachers). The design of these lecture classes and the data collection in the classes have been very similar for several years. A similar procedure was described in two previous studies with previous cohorts of students and other research questions (Bosch et al., 2021; Bosch & Spinath, 2023). In this study we assessed data from students attending the lecture classes in the winter semester 2021/2022. The lecture classes are posited early in the curriculum for both cohorts and aim at equipping students with basic knowledge on educational psychological topics for example determinants of learning success. The lecture classes were both held by the same instructor, had similar content and the students were offered the same learning activities at the same times.

Over the course of the semester, students could attend up to 13 online lecture sessions and review the lecture slides and notes, that were provided online for download, directly after the session (also up to 13 times). Additionally, students could participate in optional learning activities after each lecture session. They could submit essays via an eLearning platform answering deepening questions similar to those asked in the final exam and reflecting on the content of the last lecture session eight times. On these essays, students received individual written feedback from student tutors. After five lecture sessions (every three weeks), students could participate in online knowledge tests. The first test took place directly after the first lecture session and from that on approximately every three weeks. Within these tests, students were quizzed on the content from the last lecture sessions using confidence-weighted true/false items (Dutke & Barenberg, 2015), a type of items that was also used in the final exam of the course. Each item consisted of a statement for which students had to indicate whether it was true or false while indicating their confidence in their answer by choosing one of the four options true!, true?, false?, false!. Within one week after the test, students received feedback on their performance (number of correctly answered items). For knowledge performance scores, confidence was irrelevant.

From knowledge tests we computed students' performance, correct confidence, and accuracy of self-assessment of their knowledge. After each knowledge test, students were asked to answer questions in order to assess further study variables. At the beginning of the semester (T1), we assessed students' demographic characteristics as well as learning prerequisites (motivation and prior achievement). At each measurement occasion we assessed learning activity use. Besides the variables for the current study, we assessed further scales within the online surveys to illustrate specific contents of the lecture or to evaluate the lecture class. These variables, however were not used for the present study.

Participation in surveys was voluntary and anonymous (students generated anonymous codes for longitudinal monitoring). The procedure was in accordance with human subjects principles and procedures in Germany, which do not require formal review for anonymous survey studies like the one presented here by default. The survey was an integral part of a lecture course. Participation was anonymous and voluntary, so that individual non-participation could not be tracked. Students were informed that they would have no disadvantage if they did not participate, although participation was strongly recommended as a learning opportunity. Further, students were fully informed about content and aims of the surveys and gave their consent to the use of their data for scientific purposes.

4.2 Sample

At T1, $N = 331$ students participated in the online survey, of those $n = 138$ were psychology undergraduates and $n = 192$ preservice teachers. Psychology undergraduate students had a significantly better high school grade point average (HSGPA; prior achievement) than preservice teachers ($t(325) = 6.06$, $p < .01$, $d = 0.68$) but did not differ regarding any other variable at T1 (all *p*s > .05). Because of the parallel nature of the lecture classes' content, design, and data collection in both cohorts, all students who participated in the surveys were treated as one sample. Participants were on average $M = 21.82$ years old ($SD = 3.96$). 71.60 % identified as female, 27.50 % as male and 0.90 % as non-binary. At T2, $n = 236$ (71.3 %) students participated, at T3, $n = 199$ (60.1 %), at T4, $n = 177$ (53.5 %) and at T5, $n = 166$ (50.2 %). Missing data analyses with independent samples *t*-tests comparing the students who participated at T5 and from which we obtained dependent measures with the ones who dropped out before revealed no significant differences between groups at T1 ($t(329) = -1.17$–0.70, all $p > .05$) except for prior achievement. Students' who dropped out before T5 had a significantly worse high school grade point average than those who stayed in the sample, $t(329) = 5.56$, $p < .001$, $d = 0.61$.

4.3 Measures

To assess students' *learning activity use* (T2–T5) we asked students four times how much they had used each activity (lecture attendance, review of slides and notes, submission of essays, and participation in knowledge tests) so far. The values for possible activity use differed between time points because the lecture class continued. For example, at T3, students could have attended a maximum of 7 lecture sessions (T5: max. 13) and reviewed lecture slides a maximum of 7 times (T5: max. 13), they could have submitted a maximum of 4 essays so far

(T5: max.8) and participated in a maximum of 3 knowledge-tests (T5: max. 5). To determine the minimum use of learning activities for each student, only the learning activity use reported on the most recent measurement occasion was considered. Possible values ranged from 0-13 for lecture attendance and review of slides, from 0-8 for submission of essays and from 0-5 for participation in knowledge-tests. For analyses regarding RQ3 a mean over all learning activities was computed for each student.

To assess *prior achievement* (T1), we asked students to indicate their HSGPA. In Germany, the HSGPA ranges from 1 to 4, with lower values representing better grades. For the analyses in this paper, we recoded the HSGPA so that higher values represented better grades.

Prior knowledge (T1) regarding educational psychology was assessed with the first knowledge test at the beginning of the semester, that consisted of 15 items. An example item was "Dyslexia usually comes along with a lower IQ". Students had to indicate whether they thought the statement was true or false and indicated their confidence in their answer. If their answer was correct, students received 1 point (irrespective of their confidence). In order to compare scores from knowledge tests over time we calculated the percentage of correctly answered items.

Motivation for educational psychology (T1) was assessed with six items measuring the *expectancy* component (adapted for the "educational psychology" context instead of "school" from the Scales for measuring academic ability self-concept; SESSKO; Schöne et al., 2002). We rephrased the items to assess future-oriented and subject-specific expectancy rather than students' more stable beliefs about their ability (see also Eccles & Wigfield, 2020). An example item was "I think that I will do well in educational psychology tasks". Nine items were assessed to measure *subjective values* (adapted from the Scale for Assessing Subjective School-related Values; SESSW; Steinmayr & Spinath, 2010). Students were asked how much each statement applied to them personally. All items could be answered on a scale from 1 (*not at all*) to 5 (*very* *much*). For expectancy as well as values, a mean was computed and used in the analyses. Cronbach's α = .85 for expectancy and α = .87 for values.

Knowledge at the end of the semester (T5) was assessed with the last knowledge test consisting of 48 items for preservice teachers and 44 items for undergraduate psychology students. Preservice teachers received four more items because they received one more lecture session on a teacher-specific topic in this semester. Apart from that, the knowledge test contained four items for each lecture session in both cohorts. In order to make results comparable we calculated the percentage of correctly answered items for each student. The last knowledge tests included items that covered content of the entire semester, i.e., determinants of learning success, special learning needs, and research on educational systems among others. Therefore, the knowledge tests cannot be considered homogenous one-dimensional tests. Cronbach's α consequently indicated low values with α = 0.74 (psychology undergraduates) and α = 0.79 (preservice teachers; see also Händel et al. 2020 for a similar approach). The test scores of the prior knowledge test and the last knowledge test were significantly related with each other ($r = 0.23$, $p < .05$).

Students' *correct confidence in their own knowledge* (T1 and T5) was assessed computing an indicator that represents the match of confidence and correctness (CC) out of the knowledge test results. Using the confidence-weighted true-false items (Dutke & Barenberg, 2015) in knowledge tests, we obtained not only the number of correct answers but also whether these answers are given with confidence (*false! true!*) or uncertainty (f*alse? true?*). CC was calculated by dividing the sum of correct confident answers and incorrect uncertain answers by the sum of all items. CC would equal 1 (0) if a student gave all correct answers with confidence (uncertainty) and all incorrect answers with uncertainty (confidence; see also Barenberg & Dutke, 2019 for calculation of confidence scores).

To obtain students' *accuracy of self-assessment* (T1-T5), we asked students after each knowledge test how many items they thought they answered correctly. At T1-T4, the knowledge-test contained 15 items, therefore, also the self-assessment could range from 0 (0 Items) to 15 (15 Items). At T5, students' self-assessments could range from 1 (*0-10 Items*), to 5 (*41-all Items*). In order to obtain individual values for accuracy of self-assessment at T1 and T5 we categorized performance scores as well as self-assessments to a scale from 1 (lowest 20 %) to 5 (top 20 %) and computed absolute difference scores (accuracy as discrepancy). These scores were then recoded such that larger values represent greater accuracy. To assess the change accuracy of self-assessment over time, Spearman's rank correlations between actual performance in the knowledge tests and estimated score were calculated for each measurement occasion (accuracy in terms of rank order).

4.4 Data analysis

All data was analyzed using IBM SPSS Statistics (version 27) and Mplus 8.0 (Muthén & Muthén, 1998-2017). To address RQ1 we first inspected whether knowledge and confidence at the end of the semester was significantly better than at the beginning of the semester with *t*-tests for dependent samples (H1a and H1b). To investigate whether students' accuracy of self-assessment increased over the course of the semester (H1c), spearman's rank correlation coefficients were inspected and compared using Fisher's *z*-transformation and confidence intervals.

To investigate RQ2 we computed a multivariate regression analysis in Mplus. We entered learning prerequisites as independent variables in a first model (prior knowledge, prior achievement, and motivation) and mean use of learning activities in a second one and evaluated the increase in explained variance.

Addressing RQ3, we used latent profile analysis (LPA) to describe possible patterns of learning activity use in a person-centered way. We computed LPAs in Mplus with one up to six latent classes, the four different learning activities serving as indicators. To facilitate interpretation of the resulting estimates of the learning activities (indicators) in each class, all learning activities were grand mean centered. Consequently, all learning activities have a meaningful zero point, the average use of the activity in the sample so that negative (positive) values indicate low (high) activity use in comparison to the sample's average. A crucial question in grouping techniques is how many classes should be assumed. To decide on the number of latent classes we compared models with one up to six classes considering several statistical indicators as well as interpretability (Muthén & Muthén, 2000; Shanahan et al., 2013; Stringaris et al., 2013; see also Weller et al., 2020 for a guide to best practice). The Bayesian information criterion (BIC) rewards parsimony and indicates overall model fit (see Nylund et al., 2007). The Bootstrap likelihood ratio test (BLRT; McLachlan & Peel, 2000) provides a *p* value which indicates if a model with *k* classes is significantly better than a model with *k*-1 classes. Further, additional classification diagnostics were obtained. The average latent class posterior probability indicates the model's accuracy in assigning people to a class given the scores on indicator variables. Probabilities greater than .90 are ideal (Muthén & Muthén, 2000). The model's accuracy in defining classes is depicted in an entropy value. There is no agreed cutoff criterion for this value, however, entropy values near one indicate high accuracy. There are further recommendations that class solutions should not contain classes with less than 5 % of the sample or 50 cases (Shanahan et al., 2013). However, if a small class makes conceptual sense and model fit statistics support the class solution, also small classes can be accepted (Weller et al., 2020). Next to these criteria, interpretability of the class solution was evaluated.

To exploratively investigate the benefit of several learning activity patterns (RQ4) we added covariates to the final model of the latent profile analysis to investigate relationships between classes and outcomes. We used the BCH method (Asparouhov &Muthén, 2014) to test for differences in interesting outcomes between latent classes: after deciding on the number of latent classes we entered the outcomes of interest as auxiliary variables to the model and compared the outcomes with the Wald Test of Mean Equality for Potential Latent Class Predictors in Mixture Modeling (Asparouhouv & Muthén, 2007). This method works well when the entropy is large (Asparouhov &Muthén, 2014) and is considered the current best practice (Nylund-Gibson & Choi, 2018).

Further, associations between single activities and all outcomes were inspected exploratively evaluating bivariate associations.

5. Results

Descriptive data for study variables as well as bivariate correlation coefficients are depicted in table 1. Students started the semester with descriptively high motivation (both means above the scale mean) and little prior knowledge. The mean score achieved in the first knowledge test only slightly exceeded the probability of guessing (50 %) descriptively. Correct confidence at T1 indicates that correctness of answers and confidence converged in 51 % of students' answers.

Table 1

Means, standard deviations and bivariate correlations of study variables.

	n	*M* (*SD*)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1) Prior achievement (HSGPA)[a]	331	2.06 (0.71)									
(2) Prior motivation: Values (T1)	331	4.03 (0.62)	0.08								
(3) Prior motivation: Expectancy (T1)	331	3.24 (0.63)	0.18**	0.43**							
(4) Prior knowledge (T1) [b]	331	52.51 (12.59)	0.09	–0.02	–0.02						
(5) Correct confidence (T1)	331	0.51 (0.13)	–0.09	–0.11*	–0.00	0.11					
(6) Accuracy (T1) [c]	331	4.28 (0.67)	–0.01	0.09	–0.01	0.09	-0.03				
(7) Knowledge (T5) [b]	166	80.18 (12.24)	0.53**	0.15	0.17*	0.23**	0.03	–0.01			
(8) Correct confidence (T5)	166	0.70 (0.16)	0.26**	0.23**	0.18*	–0.00	0.08	–0.00	0.52**		
(9) Accuracy (T5) [c]	166	4.33 (0.68)	0.12	0.03	0.12	0.11	–0.13	–0.02	–0.04	0.08	
(10) Mean use of learning activities	285	5.44 (2.52)	0.25**	0.06	0.01	–0.02	–0.08	–0.01	0.46**	0.34**	0.17*

Note. [a] HSGPA: high school grade point average; larger numbers represented better grades. [b] score in knowledge test in percent. [c] absolute difference score of categorized variables (accuracy as discrepancy), recoded, larger values represent greater accuracy; * $p < .05$. ** $p < .01$.

5.1 Change of students' knowledge, correct confidence, and accuracy

At the end of the semester, students performed significantly better on the knowledge test than they did on the prior knowledge test at the beginning, $t(148) = 22.59, p < .001, d = 1.81$ (*H1a*) and were more confident in their own knowledge while being uncertain about unknown content (correct confidence) at the end of the semester than in the beginning, $t(148) = 10.22$, $p < .001$, $d = 0.84$ (*H1b*).

To investigate if students' accuracy of self-assessment in terms of rank order increased over time (*H1c*), we inspected Spearman's rank correlation between actual and estimated score in the knowledge test for each measurement occasion. Results are depicted in Table 2.

Table 2

H1c: Accuracy in terms of rank order. Correlations between actual and estimated scores across measurement occasions.

	T1	T2	T3	T4	T5
ρ	0.12*	0.26**	0.37**	0.56**	0.56**
Z	0.12	0.26	0.39	0.63	0.63
99 % CI	−0.02–0.27	0.08–0.41	0.19–0.52	0.40–0.68	0.40–0.69

Note. Estimated scores were categorized to range from 1-5, actual scores in percent. Z was obtained using a Fisher's Z transformation. 99 % CIs were calculated according to Fieller et al. (1957)
* $p < .05$. ** $p < .01$.

On a descriptive basis, correlation coefficients between actual and estimated scores increased over time from T1 to T5. Examination of the 99 % confidence intervals of the coefficients shows that the correlation coefficients from the second half of the semester (T4 and T5) were both significantly larger than the coefficient at the beginning of the semester (T1). The accuracy of self-assessment significantly improved over time, and students became more precise in estimating their level of knowledge within the sample.

5.2 Learning Activity Use Explaining Learning Outcomes

A multivariate regression analysis was computed to explain students' knowledge, correct confidence, and accuracy of self-assessment at the end of the semester (RQ2). After considering students' learning prerequisites in an initial model, including learning activity use to the model resulted in a significant increase in explained variance for all learning outcomes ($\Delta R^2 = .16$ for knowledge; $\Delta R^2 = .05$ for correct confidence; $\Delta R^2 = .04$ for accuracy of self-assessment; all ps < .01). Students who used more learning activities performed better in the last knowledge test and had a greater awareness of their own content specific strengths and weaknesses as well as their overall knowledge-level at the end of the semester. The results of the multivariate regression are shown in figure 1.

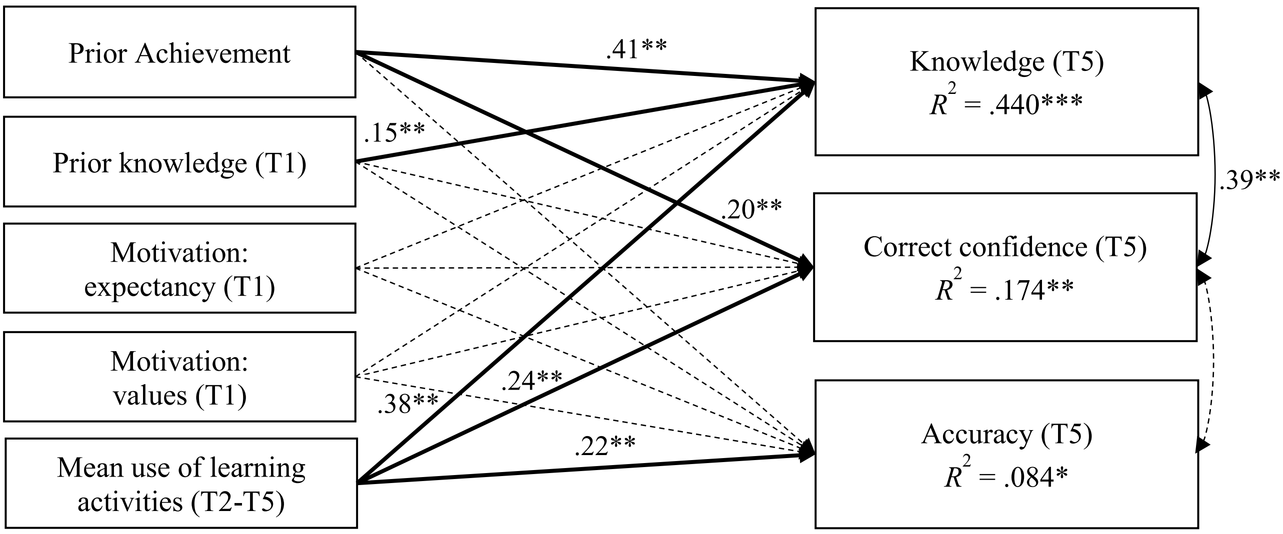


Fig. 1. Multivariate Regression Model Explaining Knowledge, Correct Confidence, and Accuracy at the End of the Semester (H2a-c).
Note. $N = 166$. The multivariate regression model shows associations between learning activity use and learning outcomes, controlling for learning prerequisites (prior achievement, prior knowledge, and motivation). Statistics presented are standardized regression coefficients. Dotted lines represent nonsignificant relations.
* $p < .05$. ** $p < .01$. *** $p < .001$

5.3 Use of Evidence-Based Learning Activities

Given the variety of learning activities we were interested in their use by students in more detail. Therefore, we aimed at discovering patterns of learning activity use. We obtained learning activity use of $N = 285$ students who had participated at least in two measurement occasions (activity use was assessed from T2 on) and indicated their use of all learning activities. Data from these 285 students were included in the latent profile analysis. Table 3 shows the statistical indicators of model fit and accuracy for all estimated models (one class up to six classes).

Table 3

Evaluating class solutions.

Models	Model fit criteria			Diagnostic criteria				Inter-pretability
	LL	BIC	BLRT	Smallest class count (*n*)	Smallest class size (%)	Entropy	ALCPP	
1 Class	–2609	5263.71	-	285	100	1	1	good
2 Classes	–2274	4622.57	*p* <.01	130	45.6	0.93	0.98	good
3 Classes	–2233	4569.73	*p* <.01	11	3.8	0.96	0.97	good
4 Classes	–2167	4465.19	*p* <.01	12	4.2	0.93	0.95	good
5 Classes	**–2143**	**4434.07**	***p* <.01**	**15**	**5.2**	**0.94**	**0.94**	**good**
6 Classes*	–2118	4392.23	*p* <.01	13	4.6	0.94	0.96	difficult

Note. $N = 285$. LL = log-likelihood; BIC = Bayesian information criterion; BLRT = Bootstrap Likelihood Ratio Test; ALCPP = average latent class posterior probability. Bold values represent the best fitting model. The model became unstable with the 6-class model. * Best Log Likelihood value not replicated.

The model became unstable with the 6-class model. This might indicate model misidentification, so we chose to inspect models with up to five latent classes (see also Marsh et al., 2009). All model solutions yielded good classification accuracy (entropy > .90). Based on the smallest BIC, we decided to continue with the 5-class solution for further analyses. Figure 2 shows how subgroups of students used each learning activity on average in comparison to the entire sample.

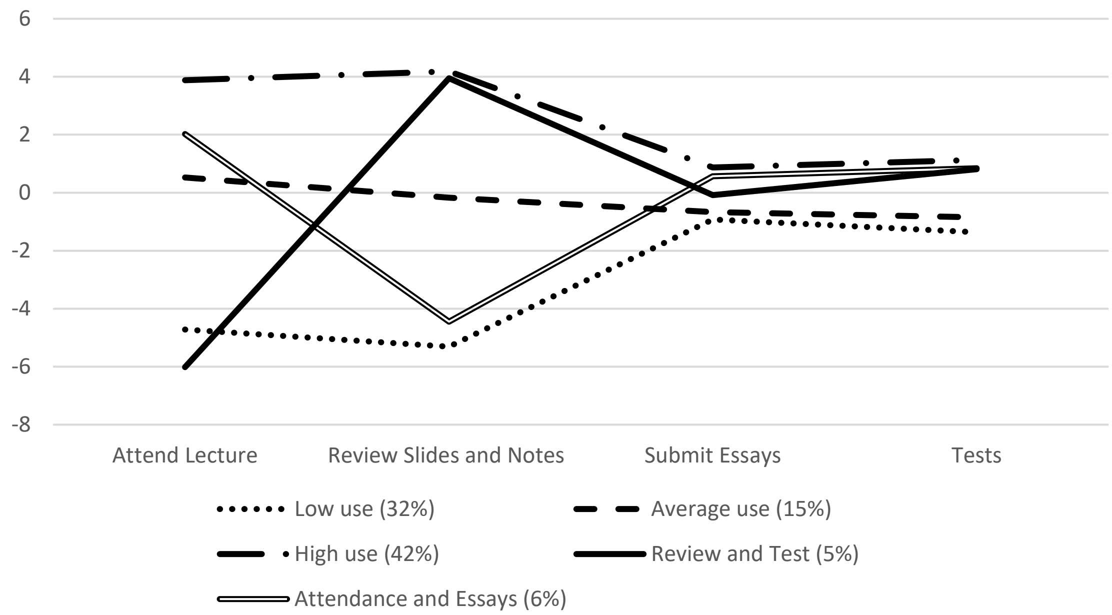


Fig. 2. Learning Activity Use in Student Subgroups.
Note. Possible use of each activity: Attendance 0–13, Review Notes 0–13, Submit Essays 0–8, Knowledge–Tests 1–5. All variables are grand mean centered.

The greatest differences between the groups were seen in lecture attendance and review of lecture slides and notes. In one group students used all activities below average (32 % of students; low use group), another group used each activity approximately on average (15 % of students; average use group) and a third group used all activities above average (42 % of students; high use group). So far, there are only quantitative differences in activity use between groups. The last two groups were small and differed in their pattern of activity use. The first group of those focused on review and tests and were therefore named accordingly (5 % of students). The last group showed an inverse pattern of activity use. These students focused on lecture attendance and were among the ones who submitted the most essays (6 % of students; attendance and essay group). The centered use of each activity in each group is depicted in table 4.

Table 4

Activity use in different subgroups of students in comparison to the sample's mean.

	Learning Activities			
	Lecture Attendance	Review of Slides and Notes	Submission of Essays	Participation in knowledge-tests
Group	*M*	*M*	*M*	*M*
Low use	–4.72	–5.31	–0.92	–1.37
Average use	0.53	–0.17	–0.67	–0.85
High use	3.89	4.19	0.87	1.14
Review and tests	–6.02	3.95	–0.08	0.81
Attendance and essays	2.02	–4.45	0.57	0.85

Note. $N = 285$. All data grand mean-centered.

5.4 Differentiated Associations between Learning Activity Use and Learning Outcomes

Although the two groups focusing on specific learning activities were small, we conducted an exploratory comparison of outcomes among these student groups, using a Bonferroni-corrected α < .005. The estimated mean outcomes for each subgroup are depicted in table 5 as well as the result of the overall Wald-Test for mean differences (X^2).

Table 5

Learning outcomes (t5) in subgroups according to students' learning activity profile.

	Knowledge *M (SE)*	Correct Confidence *M (SE)*	Accuracy *M (SE)*
X^2 (df = 4)	45.43 **	24.83 **	4.78
Low use	69.01 (3.25) ^{bc}	0.59 (0.03) ^{b}	4.08 (0.22) ^{a}
Average use	75.69 (3.41) ^{abc}	0.67 (0.04) ^{ab}	4.18 (0.22) ^{a}
High use	83.68 (1.01) ^{a}	0.74 (0.02) ^{a}	4.42 (0.06) ^{a}
Review and tests	77.97 (2.37) ^{ab}	0.61 (0.05) ^{ab}	4.20 (0.18) ^{a}
Attendance and essays	67.03 (2.96) ^{c}	0.60 (0.04) ^{b}	4.09 (0.22) ^{a}

Note. Knowledge: score in knowledge test in percent Accuracy: Difference between actual and estimated scores in the final knowledge test (accuracy as discrepancy), recoded, possible values 0–5, larger values indicate greater accuracy. Significantly differing scores within one outcome ($p < .005$) are marked with different superscripts. ** $p < .01$.

The Wald-Test with the dependent variable knowledge indicated that students from the high use group performed better in the knowledge test at the end of the semester than students from the low use group ($X^2 = 18.58$, $p < .005$) and also better than students who focused on attendance and essays ($X^2 = 28.51$, $p < .005$). Further, the group who focused on review and tests outperformed the group who focused on attendance and essays ($X^2 = 8.34$, $p < .005$). Confidence and correctness of the answers in the final knowledge test converged more often among students from the high use group than among students from the low use group ($X^2 = 16.02$, $p < .005$) and more often than among students who focused on attendance and essays ($X^2 = 11.95$, $p < .005$). There were no differences between groups in accuracy of self-assessment.

To understand the use of single activities and associations with learning outcomes in more detail exploratively we further inspected their overall use in the sample as well as bivariate associations between single learning activities and learning outcomes (table 6).

Table 6

Bivariate correlations (Spearman's rho) for single learning activities and outcomes.

	M *(SD)*	Attend Lecture	Review slides & notes	Submit essays	Knowledge	Correct Confidence	Accuracy
Attend Lecture [a]	8.44 (4.27)	-			0.37**	0.35**	0.11
Review slides and notes [a]	7.89 (4.65)	0.71**			0.42**	0.29**	0.15
Submit essays [b]	1.84 (1.18)	0.64**	0.63**		0.39**	0.23**	0.16*
Participate in knowledge tests [c]	3.58 (1.39)	0.63**	0.68**	0.67**	0.18*	0.16*	0.12

Note. [a] possible values 0–13. [b] possible values 0–8. [c] possible values 0–5. Accuracy: Difference between actual and estimated scores in the final knowledge test (accuracy as discrepancy; categorized and recoded), larger values indicate greater accuracy. * $p < .05$. ** $p < .01$.

Students attended and reviewed approximately 8 lecture sessions on average (this corresponds to 60.5 % of the maximum offered sessions). They submitted on average just under 2 essays (approximately 25 % of essays that were possible) and participated in more than 3 out of 5 possible knowledge tests (approximately 60 %). The individual activities were significantly correlated with each other.

While accuracy of self-assessment was related to submission of essays only (students who submitted more essays reported slightly more accurate estimates of their knowledge levels), knowledge as well as correct confidence were significantly associated with all learning activities with small to medium size coefficients.

6. Discussion

We investigated the acquisition of knowledge, correct confidence, and accuracy of self-assessment in a real learning situation by examining differences between the beginning and the end of a semester and investigating possible associations with students' use of evidence-based learning activities. In the following, we will discuss our main findings, limitations, and directions for future research.

6.1 Improved learning outcomes and associations with learning activities

Students started the semester with high motivation but relatively little knowledge, little correct confidence and their self-assessment of their own knowledge was correlated with actual knowledge scores only with a small coefficient (accuracy in terms of rank order). At the end of the semester, students had improved in all learning outcomes. One important explanation for these improvements over the course of the semester were evidence-based learning activities. The more students used the optional activities, the better their learning outcomes were beyond individual learning prerequisites (prior knowledge, prior achievement and motivation).

Especially in knowledge, the use of learning activities explained substantial additional variance (16 %). This finding replicates prior findings with the same evidence-based learning activities in other cohorts of students (Bosch et al., 2021) and emphasizes the importance of students' active engagement and behavior in order to succeed in higher education. While the use of learning activities explained acquisition of knowledge with a medium to large effect, associations with correct confidence and accuracy of self-assessment were rather small. Other studies that investigated correct confidence in students yielded improvements with small to medium effect sizes (Barenberg & Dutke, 2019; Carvalho et al., 2022; Naujoks et al., 2022). Most studies investigated effects of practice testing only, oftentimes neglecting the use of other learning activities (see Naujoks et al., 2022). Surprisingly, in bivariate associations of single learning activities with learning outcomes, participation in knowledge tests yielded rather small coefficients in comparison to other learning activities (see table 6). Although this result is only exploratory, it is still noteworthy that all learning activities were at least equally associated with learning outcomes as knowledge tests.

Looking at accuracy of self-assessment, the absolute difference of achieved and estimated scores at the end of the semester was only associated with the use of learning activities with a small coefficient. One possible methodological explanation could be that the difference-score for accuracy was not reliable enough to reveal associations with learning activities because difference-scores can be unreliable (Lord, 1958). In our case the difference score could not differentiate between over- and underestimation of students' own knowledge. On the other hand, examining the rank correlations, it is evident that self-assessment and actual knowledge levels were becoming more similar over time. This could also have been due to the fact that students were repeatedly asked to self-assess their knowledge in this format (see also Händel et al., 2020). What other factors might contribute to an increase in accuracy of self-assessment over the course of a semester would be of great interest for future studies.

Models of self-regulated learning suggest that correct confidence and an accurate assessment of one's own knowledge level are crucial to direct self-regulated learning (see 2.; Schmitz & Wiese, 2006; Winne, 1996; Zimmerman, 2008). In our sample, indeed, correct confidence was associated with more knowledge at the end of the semester. Maybe, more correct confidence had led to a better allocation of resources for learning unknown content. This mechanism could be further investigated in future research. In contrast, accuracy of self-assessment was not related to knowledge in our study. Models of self-regulated learning assume that accurate assessment of one's knowledge level is critical to efficiently direct learning. However, efficient learning may not necessarily mean *more* learning for all students.

6.2 Profiles of learning activity use

It is important to assess in detail what students do over the course of one semester to understand how students actually study in higher education (see also Blasiman et al., 2017). Students in our study used all activities less than would have been possible. On average they preferred knowledge tests, lecture attendance and review of lecture notes over submission of essays. Further, students in our study used optional evidence-based learning activities, at least to a small extent, in different combinations. We identified five different patterns of activity use. We had no hypotheses about the associations of specific activity use patterns and learning outcomes. However, some activities are maybe more appealing for some students than for others. It would have been plausible that students might focus on individually most appropriate learning activities. In our data, individual learning activities were highly intercorrelated. Additionally, in the profile analysis, the majority of students (89 %) could be grouped into subgroups distinguished mainly by their overall engagement in learning activities. This observation might imply that, despite their variations, the use of all learning activities could serve as an indicator of overall student engagement. In our specific case, latent profile analyses did not uncover large subgroups of students who employed distinct combinations of activities. Our data suggest a "the more, the better" model. Students in the profile with above-average use of all activities outperformed several other subgroups of students in knowledge acquisition and correct confidence confirming the results from the multivariate regression analysis. In line with these results, use of all activities was associated with acquisition of knowledge as well as correct confidence on a bivariate basis. This finding complements many studies that have largely focused on the benefits of practice testing mainly (e.g., Barenberg & Dutke, 2019; Naujoks et al., 2022; Yang et al., 2021). Our results confirm the usefulness of all activities, strengthening the evidence for well-established learning activities such as practice testing and distributed practice (e.g., Dunlosky et al., 2013) but also regular class attendance and review of lecture material in higher education (Credé et al., 2010). Future studies examining the benefits of evidence-based learning activities such as testing might benefit from considering other possible learning activities as well. This would help to understand more precisely which learning activities contribute to learning success and how much. Further, future research might look at different student subgroups according to their learning activity use over time. Continuous and distributed learning is recommended (e.g., Dunlosky et al., 2013; Naujoks et al., 2022) and future research could examine different time-patterns of activity use.

The current study expands the state of research in multiple ways. First, it looks at multiple outcomes simultaneously, expanding the evidence on the usefulness of learning activities for different aspects of student learning, i.e., for the acquisition of knowledge but also correct confidence, and an accurate self-assessment. Second, the study assesses the use of different practice testing activities (as recommended by Carvalho et al., 2022) as well as other optional learning activities and draws a more realistic picture of what students do in lecture classes that might contribute to learning success (see also Blasiman et al., 2017; Naujoks et al., 2022).

6.3 Limitations and directions for future research

Of course, there are limitations to consider when interpreting the results of the current study. First of all, due to the non-experimental study design, we cannot draw causal conclusions. However, in order to obtain a realistic picture of students' learning behavior in a lecture and to account for different learning activities, we decided to capture students' use of activities when activities are offered rather than manipulating activity use in an experimental design. However, by using a longitudinal design and controlling for important predictors of learning success, we can assume that the use of learning activities over time contributes to student learning success. Future research could employ an experimental design, considering multiple learning activities simultaneously. We believe it is realistic for students to use different learning activities in self-directed learning situations. How students choose learning activities and whether this choice meets individual learning needs as well as individual goals (e.g., just passing an exam rather than getting a good grade, perhaps to have more resources to pursue other goals) should be further investigated to understand self-regulated learning in all its facets.

Another limitation concerns the sample. We focused on students who attended a specific lecture class. While this has the advantage of keeping many relevant factors constant (content of the lecture and knowledge tests, instructor, learning activities), it also limits the generalizability of the results. The results are consistent with previous research on the benefits of evidence-based learning activities for different learning outcomes (e.g., Barenberg & Dutke, 2019; Bosch et al., 2021; Dunlosky et al., 2013; Naujoks et al., 2022). However, it would be desirable for future research to consider additional samples from a variety learning situations to support the importance of providing evidence-based learning activities in higher education regardless of the field of study. Future research could also aim to objectively measure students' learning activity use, which was not possible in the current study. Within our study design this would have meant revealing students' identity and we wanted to collect anonymous data. We cannot rule out that students' report of learning activity use was not accurate (see, e.g., Blasiman et al., 2017). However, by frequently asking students about their specific learning activity use and disentangling it from their grade, we believe that students were able to correctly report their learning activity use. Another limitation in the context of our sample is dropout. We had relatively high dropout rates, so our results may be biased in favor of more motivated and committed students and we cannot draw conclusions for students who dropped out. However, the dropout analyses showed that there were no differences in motivation in the beginning. The differences in prior achievement may suggest that students with better learning habits (and consequentially better high school grade point averages) continue their engagement, participate in studies as well as in optional learning activities, and consequently learn more. It would be of interest for future research to identify factors that help students with lower prior achievement adopting beneficial study habits and engaging in learning activities.

In addition, the latent profile analysis mainly confirmed the regression analysis' results that more activity use is associated with better learning outcomes. Subgroups of students applying a differentiated activity use pattern were small limiting the ability to compare results between these groups. In this study the latent profile analysis failed to add significant results regarding specific activity patterns and their usefulness for learning outcomes. It would be important to repeat the latent profile analyses with a different and likely larger sample to replicate the class solution or to draw the conclusion that there are no interindividual differing patterns of activity use. Our sample size was slightly smaller than 300, a minimum sample size recommended, for example, by Nylund-Gibson and Choi (2018) for latent class analyses and our model became unstable with the six-class solution. It cannot be ruled out that with a larger sample, a model with more classes would represent the data even better.

6.4 Implications and conclusion

Holding these limitations in mind, the current study extends existing research in several ways. By examining various learning activities and their associations with multiple indicators of learning success in a realistic learning situation, we reinforced the importance of providing evidence-based activities in higher education. Greater use of learning activities was positively associated with all desirable learning outcomes. Furthermore, to address the complex situation in the context of self-regulated learning, the use of learning activities was examined using a person-centered approach. The results of both the regression analyses explaining the learning outcomes and the comparison of the different latent profiles support a "the more, the better" model. Future research should consider various student learning activities and goals to understand the complex learning process in higher education and what factors contribute to student success.

References

Asparouhouv, T., & Muthén, B. (2007). Wald test of mean equality for potential latent class predictors in mixture modeling. Retrieved from http://www.statmodel.com/download/MeanTest1.pdf

Asparouhov, T., & Muthén, B. (2014). Auxiliary variables in mixture modeling: A 3-step approach using M*plus*. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(3), 329-341. https://doi.org/10.1080/10705511.2014.915181

Barenberg, J., & Dutke, S. (2019). Testing and metacognition: retrieval practise effects on metacognitive monitoring in learning from text. *Memory*, *27*(3), 269-279. https://doi.org/10.1080/09658211.2018.1506481

Barenberg, J., Roeder, U. R., & Dutke, S. (2018). Students' temporal distributing of learning activities in psychology courses: Factors of influence and effects on the metacognitive learning outcome. *Psychology Learning & Teaching*, *17*(3), 257-271. https://doi.org/10.1177/1475725718769488

Bae, C. L., Therriault, D. J., & Redifer, J. L. (2019). Investigating the testing effect: Retrieval as a characteristic of effective study strategies. *Learning and Instruction*, *60*, 206–214. https://doi.org/10.1016/j.learninstruc.2017.12.008

Bensley, D. A., Rainey, C., Lilienfeld, S. O., & Kuehne, S. (2015). What do psychology students know about what they know in psychology?. *Scholarship of Teaching and Learning in Psychology*, *1*(4), 283-297. http://dx.doi.org/10.1037/stl0000035

Blasiman, R. N., Dunlosky, J., & Rawson, K. A. (2017). The what, how much, and when of study strategies: Comparing intended versus actual study behaviour. *Memory*, *25*(6), 784-792. http://dx.doi.org/10.1080/09658211.2016.1221974

Bosch, E., Seifried, E., & Spinath, B. (2021). What successful students do: Evidence-based learning activities matter for students' performance in higher education beyond prior knowledge, motivation, and prior achievement. *Learning and Individual Differences*, *91*, 102056. https://doi.org/10.1016/j.lindif.2021.102056

Bosch, E., & Spinath, B. (2023). Students' motivation in an online and a face-to-face semester. A comparison of initial level, development, and use of learning activities. *Zeitschrift für Psychologie, 231*(2), 93-102. https://doi.org/10.1027/2151-2604/a000519

Boser, J., Scherer, S., Kuchta, K., Wenzel, S. F. C., & Horz, H. (2017). Empirically founded teaching in psychology–An example for the combination of evidence-based teaching and the scholarship of teaching and learning. *Psychology Learning & Teaching*, *16*(2), 261-275. https://doi.org/10.1177/1475725716686452

Brühwiler, C., & Blatchford, P. (2011). Effects of class size and adaptive teaching competency on classroom processes and academic outcome. *Learning and Instruction*, *21*(1), 95-108. https://doi.org/10.1016/j.learninstruc.2009.11.004

Carrier, L. M. (2003). College students' choices of study strategies. *Perceptual & Motor Skills, 96*(1), 54–56, https://doi.org/10.2466/PMS.96.1.54-56

Carvalho, P. F., McLaughlin, E. A., & Koedinger, K. R. (2022). Varied practice testing is associated with better learning outcomes in self-regulated online learning. *Journal of Educational Psychology, 14*(8), 1723-1742. https://doi.org/10.1037/edu0000754

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, *132*(3), 354. https://doi.org/10.1037/0033-2909.132.3.354

Credé, M., & Kuncel, N. R. (2008). Study habits, skills, and attitudes: The third pillar supporting collegiate academic performance. *Perspectives on Psychological Science, 3*(6), 425–453. https://doi.org/10.1111/j.1745-6924.2008.00089.x

Credé, M., Roch, S. G., & Kieszczynka, U. M. (2010). Class attendance in college: A meta-analytic review of the relationship of class attendance with grades and student characteristics. *Review of Educational Research*, *8*(2), 272–295. https://doi.org/10.3102/0034654310362998

Downs, S. D. (2015). Testing in the college classroom: Do testing and feedback influence grades throughout an entire semester? *Scholarship of Teaching and Learning in Psychology, 1*(2), 172–181. https://doi.org/10.1037/stl0000025

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, *14*(1), 4–58. https://doi.org/10.1177/1529100612453266

Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, *12*, 83–87. https://doi.org/0.1111/14678721.01235

Dutke, S., & Barenberg, J. (2015). Easy and informative: using confidence-weighted true–false items for knowledge tests in psychology courses. *Psychology Learning & Teaching, 14*(3), 250-259. https://doi.org/ 10.1177/1475725715605627

Eccles, J.S., & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology*, *61*, 101859. https://doi.org/10.1016/j.cedpsych.2020.101859

Fernandez, J., & Jamet, E. (2016). Extending the testing effect to self-regulated learning. *Metacognition and Learning, 12*(2), 131–156. https://doi.org/10.1007/s11409-016-9163-9

Fieller, E. C., Hartley, H. O., & Pearson, E. S. (1957). Tests for rank correlation coefficients. I. *Biometrika*, *44*(3/4), 470-481. https://doi.org/10.2307/2332878

Foster, N. L., Was, C. A., Dunlosky, J., & Isaacson, R. M. (2017). Even after thirteen class exams, students are still overconfident: The role of memory for past exam performance in student predictions. *Metacognition and Learning, 12*, 1–19. https://doi.org/10.1007/s11409-016-9158-6.

Hacker, D. J., Bol, L., Horgan, D. D., & Rakow, E. A. (2000). Test prediction and performance in a classroom context. *Journal of Educational Psychology, 92*, 160–170. doi:10.1037/0022-0663.92.1.160

Hartwig, M. K., & Dunlosky, J. (2014). The contribution of judgment scale to the unskilled-and-unaware phenomenon: How evaluating others can exaggerate over- (and under-) confidence. *Memory & Cognition, 42*, 164–173. http://dx.doi.org/10.3758/s13421-013-0351-4

Hattie, J. (2011). Which strategies best enhance teaching and learning in higher education? In D. Mashek, & E.Y. Hammer (Eds.), *Empirical research in teaching and learning: Contributions from Social Psychology* (pp.130–142). London: Blackwell Publishing.

Händel, M., & Dresel, M. (2018). Confidence in performance judgment accuracy: The unskilled and unaware effect revisited. *Metacognition and Learning*, *13*, 265-285. https://doi.org/10.1007/s11409-018-9185-6

Händel, M., Harder, B., & Dresel, M. (2020). Enhanced monitoring accuracy and test performance: Incremental effects of judgment training over and above repeated testing. *Learning and Instruction, 65*, 101245. https://doi.org/10.1016/j.learninstruc.2019.101245

Kornell, N., & Bjork, R. A. (2007). The promise and perils of self-regulated study. *Psychonomic Bulletin & Review*, *14*(2), 219–224. https://doi.org/10.3758/BF03194055

Kostons, D., van Gog, T., & Paas, F. (2012). Training self-assessment and task-selection skills: A cognitive approach to improving self-regulated learning. *Learning and Instruction, 22,* 121–132. https://doi.org/10.1016/j.learninstruc.2011.08.004.

Kriegbaum, K., Becker, N., & Spinath, B. (2018). The relative importance of intelligence and motivation as predictors of school achievement: A meta-analysis. *Educational Research Review*, *25*, 120–148. https://doi.org/10.1016/j.edurev.2018.10.001

Lord, F. M. (1958). The utilization of unreliable difference scores. *Journal of Educational Psychology, 49*(3), 150–152. https://doi.org/10.1037/h0045456

Marsh, H. W., Lüdtke, O., Trautwein, U., & Morin, A. J. (2009). Classical latent profile analysis of academic self-concept dimensions: Synergy of person-and variable-centered approaches to theoretical models of self-concept. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(2), 191-225.

Miller, T. M., & Geraci, L. (2011). Training metacognition in the classroom: The influence of incentives and feedback on exam predictions. *Metacognition and Learning*, *6*, 303-314. https://doi.org/10.1007/s11409-011-9083-7

McLachlan, G., & Peel, D. (2000). *Finite mixture models*. Wiley.

Muthén, L.K., & Muthén, B.O. (1998 –2017). *Mplus user's guide* (8th ed.). Los Angeles, CA: Author.

Muthén, B., & Muthén, L. K. (2000). Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and experimental research*, *24*(6), 882-891. https://doi.org/10.1111/j.1530-0277.2000.tb02070.x

Naujoks, N., Harder, B., & Händel, M. (2022). Testing pays off twice: Potentials of practice tests and feedback regarding exam performance and judgment accuracy. *Metacognition and Learning*, *17*, 479-498. https://doi.org/10.1007/s11409-022-09295-x

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural equation modeling: A multidisciplinary Journal*, *14*(4), 535-569. https://doi.org/10.1080/10705510701575396

Nylund-Gibson, K., & Choi, A. Y. (2018). Ten frequently asked questions about latent class analysis. *Translational Issues in Psychological Science, 4*(4), 440-461. https://doi.org/10.1037/tps0000176

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. https://doi.org/10.1037/a0026838

Rivers, M. L., Dunlosky, J., & Joynes, R. (2019). The contribution of classroom exams to formative evaluation of concept-level knowledge. *Contemporary Educational Psychology*, *59*, 101806. https://doi.org/10.1016/j.cedpsych.2019.101806.

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin*, *130*(2), 261-288. https://doi.org/10.1037/0033-2909.130.2.261

Rowland, C. A. (2014). The effect of testing versus restudy on retention: A meta-analytic review of the testing effect. *Psychological Bulletin, 140*(6), 1432–1463. http://dx.doi.org/10.1037/a0037559

Rutkiene, A. & Tandzegolskiene, I. (2015). Students' attitude towards learning methods for self-sufficiency development in higher education. In V. Lubkina & S. Usca (Eds), Society. Integration. Education. Proceedings of the International Scientific Conference (pp. 348-357). Rēzekne, Latvia: Rēzeknes Augstskola

Sackett, P. R., Kuncel, N. R., Arneson, J. J., Cooper, S. R., & Waters, S. D. (2009). Does socioeconomic status explain the relationship between admissions tests and post-secondary academic performance?. *Psychological Bulletin*, *135*(1), 1-22. https://doi.org/0.1037/a0013978

Schindler, J., & Richter, T. (2023). Text Generation Benefits Learning: a Meta-Analytic Review. *Educational Psychology Review,* **35**, 44. https://doi.org/10.1007/s10648-023-09758-w

Schneider, M., & Preckel, F. (2017). Variables associated with achievement in higher education: A systematic review of meta-analyses. *Psychological Bulletin*, *143*(6), 565-600. http://dx.doi.org/10.1037/bul0000098

Schöne, C., Dickhäuser, O., Spinath, B., & Stiensmeier-Pelster, J. (2002). *Skalen zur Erfassung des schulischen Selbstkonzepts: SESSKO* [Scales for measuring academic ability self-concept]. Hogrefe.

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. https://doi.org/10.1016/j.cedpsych.2005.02.002

Seidel, T. (2014). Angebots-Nutzungs-Modelle in der Unterrichtspsychologie. Integration von Struktur-und Prozessparadigma. *Zeitschrift für Pädagogik*, *60*(6), 850-866. URN: urn:nbn:de:0111-pedocs-146864

Shanahan, L., Copeland, W. E., Worthman, C. M., Erkanli, A., Angold, A., & Costello, E. J. (2013). Sex-differentiated changes in C-reactive protein from ages 9 to 21: The contributions of BMI and physical/sexual maturation. *Psychoneuroendocrinology*, *38*(10), 2209-2217. https://doi.org/10.1016/j.psyneuen.2013.04.010

Steinmayr, R., & Spinath, B. (2010). Konstruktion und erste Validierung einer Skala zur Erfassung subjektiver schulischer Werte (SESSW) [Construction and First Validation of a Scale for Assessing Subjective School-related Values]. *Diagnostica, 56*, 195–211. https://doi.org/10.1026/0012-1924/a000023

Stone, N. J. (2000). Exploring the relationship between calibration and self-regulated learning. *Educational Psychology Review, 12*, 437–475. http://dx.doi.org/10.1023/A:1009084430926

Stringaris, A., Maughan, B., Copeland, W. S., Costello, E. J., & Angold, A. (2013). Irritable mood as a symptom of depression in youth: Prevalence, developmental, and clinical correlates in the Great Smoky Mountains study. *Journal of the American Academy of*

Child & Adolescent Psychiatry, *52*(8), 831-840. https://doi.org/10.1016/j.jaac.2013.05.017

Thiede, K.W. (1999). The importance of monitoring and self-regulation during multitrial learning. *Psychonomic Bulletin & Review* 6, 662–667. https://doi.org/10.3758/BF03212976

Thiede, K. W., Anderson, M., & Therriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology*, *95*(1), 66-73. https://doi.org/10.1037/0022-0663.95.1.66

Thomas, R. C., & McDaniel, M. A. (2013). Testing and feedback effects on front-end control over later retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(2), 437–450. http://dx.doi.org/10.1037/a0028886

Watts, M., & Schaur, G. (2011). Teaching and assessment methods in undergraduate economics: A fourth national quinquennial survey. *The Journal of Economic Education*, *42*(3), 294-309. https://doi.org/10.3200/JECE.39.3.273-286

Weller, B. E., Bowen, N. K., & Faubert, S. J. (2020). Latent class analysis: a guide to best practice. *Journal of Black Psychology*, *46*(4), 287-311.

Winne, P. H. (1996). A metacognitive view of individual differences in self-regulated learning. *Learning and Individual Differences, 8,* 327–353. https://doi.org/10.1016/S1041-6080(96)90022-9.

Winne, P. H. (2011). A cognitive and metacognitive analysis of self-regulated learning. In B. J. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation and of learning and performance* (pp. 15–32). New York, NY: Routledge.

Yang, C., Luo, L., Vadillo, M. A., Yu, R., & Shanks, D. R. (2021). Testing (quizzing) boosts classroom learning: A systematic and meta-analytic review. *Psychological Bulletin*, *147*(4), 399-435. http://dx.doi.org/10.1037/bul0000309

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal*, *45*(1), 166–183. https://doi.org/10.3102/0002831207312909